# Separate functions of the Paf1 and Ski complexes in transcription and RNA decay

**Inauguraldissertation**

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

**Aneliya Rankova**

aus Bulgarien

Basel, 2020

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Marc Bühler

Prof. Dr. Stefanie Jonas

Basel, den 19. November 2019

Prof. Dr. Martin Spiess

Dekan

# Table of Contents

# Summary

Each stage of the mRNA life cycle, from its transcription and processing in the nucleus, to its export into the cytoplasm where it is translated and eventually degraded, is subject to elaborate control. Although commonly viewed as discrete and independent events, the different stages of gene expression are often physically and functionally linked. The first chapter of this thesis is focused on the conserved Polymerase Associated Factor 1 complex (Paf1C), which regulates multiple steps of the RNA polymerase II transcription cycle as well as events downstream of transcript synthesis such as polyadenylation of mRNAs and export of nascent transcripts. Recent studies in the fission yeast *Schizosaccharomyces pombe* uncovered an additional role of Paf1C as an antagonist of RNAi-directed transcriptional gene silencing. In Paf1C mutant strains, ectopically expressed small RNAs (siRNAs) mediate gene repression by inducing *de novo* heterochromatin formation at the target locus in the nucleus. Although euchromatic loci in wild type cells seem refractory to this repressive mechanism, the induced silent gene state in Paf1C mutants can be inherited across generations even in the absence of the original siRNA source. In mammalian somatic cells, the existence of a nuclear RNAi pathway that can similarly modulate gene expression by modifying the underlying chromatin environment is highly debatable. However, the discovery of Paf1C as a negative regulator of this process in fission yeast raised the question of whether the complex performs similar repressive functions in higher organisms. Therefore, one of the goals of my PhD project was to test if in the absence of the Paf1 complex, siRNAs can initiate *de novo* heterochromatin formation in mammals, using mouse embryonic stem cells (mESCs) as a model system. However, by examining Paf1C conditional knock out cells expressing endogenous siRNAs, I did not find any supporting evidence for the existence of a nuclear RNAi silencing mechanism that can be regulated by the mammalian Paf1 complex.

Although Paf1C is conserved from yeast to humans, the complex in higher eukaryotes has an additional subunit, Ski8/Wdr61, which is also part of the helicase Ski complex functioning in cytoplasmic RNA decay. In human cells, Ski8/Wdr61 is thought to bridge the two complexes, suggesting that Paf1C might affect other stages of RNA processing or decay in the nucleus via its association with the Ski complex. Thus, another goal of my PhD project was to elucidate the functional role of the interaction between the Paf1 and Ski complexes in mammalian cells. In contrast to previous studies, my data suggest that in mESC the two complexes are not physically or functionally linked, even though they contain a common subunit. Further investigation of the mammalian Ski complex provided additional support for independent roles of the two complexes in transcription and cytoplasmic RNA decay.

The above findings became the basis for my second PhD project, described in the second chapter of the thesis. This was part of a collaborative effort examining the two major cytoplasmic mRNA degradation

pathways in mES cells. In the cytoplasm, mRNAs can be degraded in the 5′-3′ direction by the exoribonuclease Xrn1 or in the 3′-5′ direction by the RNA exosome. The latter pathway requires the function of the Ski complex, comprising the scaffold protein Ski3/Ttc37, two copies of Ski8/Wdr61 and the RNA helicase Ski2/Skiv2l. The Ski complex is suggested to facilitate substrate passage through the exosome channel during 3′-5′ decay. RNA degradation has been extensively studied in the yeast *Saccharomyces cerevisiae*, where Xrn1 seems to be the predominant route for mRNA decay, whereas the 3′-5′ Ski-exosome pathway is thought to function redundantly with Xrn1, and contribute more significantly to RNA surveillance. In mammalian cells, the specific endogenous targets of the two pathways are poorly defined and it remains unclear if certain mRNAs can be preferentially degraded via one route, and if so, what factors could mediate such specificity. In addition, cytoplasmic RNA degradation is widely influenced by translation, consistent with recent cryo-EM structures capturing the yeast Ski complex or Xrn1 bound directly to translating ribosomes. It is currently unknown whether this physical link between the translation and RNA degradation machineries is conserved in mammals. Furthermore, it is unclear to what extent each decay pathway interacts with translation, and what factors might influence this process.

We combined crosslinking and analysis of cDNA (CRAC) with ribosome profiling in mES cells to examine the two major cytoplasmic RNA decay pathways and how they are linked to translation. Our approach allowed us to determine the direct transcriptome-wide RNA targets of Xrn1 and the Ski complex helicase Skiv2l in unperturbed cells and identify a subset of transcripts whose steady-state levels depend on the 3′-5′ pathway. Strikingly, although we found that both pathways are physically linked to translation, Skiv2l binding to RNA was exclusively dictated by ribosome occupancy and was heavily dependent on the translational status of the substrate. Our data reveal diverse triggers of RNA decay, including specific amino acid codons and RNA sequences that seem to impede ribosome elongation. We further identified a novel interaction between the Ski complex and the higher-eukaryote-specific RNA binding protein Aven. We showed that Aven and Skiv2l function closely to oppose aberrant translation, with Aven helping to prevent ribosome stalling at structured regions, while Skiv2l eliminates transcripts when these events accumulate. Interestingly, the Aven-Skiv2l pathway acts on a wide range of substrates, including mRNAs, uORFs and most surprisingly on small-ORF-containing RNAs derived from transcription of non-coding regions. As Aven is conserved from Drosophila to humans, this work uncovered a higher-eukaryote-specific pathway that coordinates cytoplasmic 3′-5′ RNA decay.

# Chapter 1

## INTRODUCTION

### 1. The Paf1 complex is a versatile regulator of PolII transcription

Transcription by RNA PolII is a complex, highly regulated process that is fundamental to the control of gene expression in eukaryotic cells. In addition to the core transcriptional machinery, multiple other protein complexes and accessory factors are required to coordinate the different phases of transcription. One of these players is the Polymerase-Associated Factor 1 complex (Paf1C), which performs multiple functions throughout the transcriptional cycle. It is a highly conserved protein complex that was first identified in *Saccharomyces cerevisiae* as a positive regulator of PolII-transcribed genes (Tomson and Arndt, 2013). *S. cerevisiae* Paf1C is composed of five subunits – Paf1, Ctr9, Cdc73, Rtf1 and Leo1, which directly associate with the transcribing polymerase from the promoter to the 3′ end formation site of genes. The complex regulates multiple events during the transcription cycle including: promoter-proximal polymerase pausing, promoting histone modifications associated with active transcription and recruitment of 3′ end processing factors necessary for proper termination of transcription (Van Oss et al., 2017). Through this broad range of regulatory functions, Paf1C has been implicated in cellular processes ranging from regulation of gene expression and DNA-damage repair to cell cycle progression, and is dysregulated in several diseases (Tomson and Arndt, 2013).

Although none of the Paf1C subunits have an enzymatic activity, the complex coordinates interactions of multiple factors with the PolII machinery and thus impacts transcription at multiple levels. Ctr9 is a tetratricopeptide repeat-containing (TPR) structural protein, Cdc73 has a C-terminal Ras-like domain, which facilitates association of the complex to chromatin (Amrich et al., 2012), and Rtf1 contains a Plus3 domain which binds single-stranded DNA (de Jong et al., 2008). Leo1 and Paf1, however, do not seem to possess any distinct structural domains (Tomson and Arndt, 2013). The Paf1 complex is not essential for viability in *S. cerevisiae*, although loss of Ctr9 or Paf1 causes growth defects (Betz et al., 2002). This is in line with observations that these two subunits are central to the overall stability of the complex (Chu et al., 2013). Unlike *S. cerevisiae*, some Paf1C components in higher eukaryotes seem to be essential (Bahrampour and Thor, 2016; Mosimann et al., 2006).

Initial evidence that Paf1C positively impacts transcription is based on the direct interaction of the complex with PollII (Shi et al.; Wade et al., 1996) and its localization to actively transcribed genes as assessed by chromatin immunoprecipitation (Krogan et al., 2002; Pokholok et al., 2002). This is partially mediated by

Cdc73, which has been shown to bind directly to PolII and facilitate recruitment of the complex (Mueller et al., 2004; Shi et al., 1997). Paf1C has genetic and physical links with elongation factors such as Spt4-Spt5 and Spt16-Pob3 (Krogan et al., 2002; Squazzo et al., 2002) and has been suggested to help bring FACT and Spt6 to actively transcribed genes (Adelman et al., 2006). Further evidence that Paf1C facilitates transcription comes from *in vitro* studies, demonstrating the ability of the human complex to stimulate PolII elongation both independently and in synergy with the elongation factor TFIIS (Kim et al., 2010).

Several recent studies describe a novel role of the Paf1 complex as a regulator of PolII promoter-proximal pausing in metazoans (Chen et al., 2015; Yu et al., 2015). Many developmentally and environmentally regulated genes harbor a paused PolII 20-60nt downstream of the transcription start site, which can be released into productive transcription in response to appropriate stimuli, thus facilitating rapid induction of gene expression (Liu et al., 2015). Loss of Paf1 in *Drosophila* and human cells leads to a redistribution of paused PolII from promoters into the gene bodies. This is accompanied by an increase in nascent and mature transcript levels, suggesting that the absence of Paf1C stimulates the release of paused PolII into productive elongation (Chen et al., 2015). According to proposed mechanistic models, depletion of Paf1C facilitates PolII CTD phosphorylation on serine 2 by SEC (Ser2P kinase super elongation complex), thereby promoting transcription into the gene body (Chen et al., 2015). Interestingly, another study using different human cell lines reports that Paf1C can have both a positive and negative impact on paused PolII, suggesting its role as a regulator of promoter-proximal pausing can be cell- and context-dependent (Yu et al., 2015).

Transcription by PolII in eukaryotes is accompanied by dynamic changes in the underlying chromatin and the Paf1 complex has also been shown to affect the deposition of several co-transcriptional histone modifications. Components of the Paf1 complex interact with the ubiquitin ligase Rad6-Bre1 to stimulate monoubiquitylation of a conserved lysine on histone H2B (H2Bub) (Ng et al., 2003; Wood et al., 2003). This mark promotes an open and accessible chromatin conformation (Fierz et al., 2011) and is a prerequisite for establishing two other histone modification, histone H3 lysine 4 (H3-K4) and lysine 79 (H3-K79) methylation, with important roles in transcription (Zhu et al., 2005b). Loss of yeast Paf1C has also been shown to reduce levels of the conserved H3 K36 methylation mark which is usually found in actively transcribed genes (Chu et al., 2007). However, this might reflect a more indirect effect, since loss of Paf1C reduces PolII CTD Ser2 phosphorylation levels, leading to destabilization of the H3K36 methyltransferase SET2 (Fuchs et al., 2012).

In addition to its effects on chromatin, PolII pausing and elongation, Paf1C is also implicated in the later stages of the transcriptional cycle and contributes to proper 3′ end formation of polyadenylated transcripts. In *S. cerevisiae* loss of Paf1C subunits is linked to shortening of the polyA tail of transcripts as well as an increase in usage of distal polyA sites, leading to production of longer transcripts (Mueller et al., 2004;

Penheiter et al., 2005). In line with these observations, PafC can physically interact with the yeast polyadenylation factor Cft1 in a PolII independent manner, thus perhaps facilitating contacts between the transcription and polyadenylation machineries (Nordick et al., 2008). Furthermore, human Cdc73 has been shown to associate with the CPSF complex (Cleavage and Polyadenylation Specificity Factors) and the cleavage stimulation factor, CstF, both of which are important for maturation of mRNA 3′ ends (Rozenblatt-Rosen et al., 2009). The ability of yeast Paf1C to promote transcript cleavage and polyadenylation is further supported by in vitro experiments (Nagaike et al., 2011). Another study in mouse myoblasts reports that loss of Paf1C leads to a global increase in the use of alternative polyA sites in introns and internal exons, thus extending the role of Paf1C in mRNA 3′ formation to mammalian cells (Yang et al., 2016).

Although Paf1C is highly conserved, some important characteristics distinguish the yeast complex from its homologs in higher eukaryotes and could potentially underlie additional roles of Paf1C in more complex organisms. For instance, as mentioned above, although loss of Ctr9 and Paf1 in *S. cerevisiae* is associated with growth defects, none of the five subunits are essential for viability (Betz et al., 2002). In contrast, some Paf1C components in higher organisms are linked to more severe phenotypes (Bahrampour and Thor, 2016; Mosimann et al., 2006). Furthermore, while *S. cerevisiae* Paf1C is highly enriched shortly after the transcription start site of genes and seems to dissociate from chromatin just upstream of the polyA site , in mouse cells Paf1C binding extends beyond the polyA site (Yang et al., 2016). Finally, the complex composition in higher eukaryotes has some intriguing differences. Rtf1 does not seem to be a core Paf1C component in organisms other than *S. cerevisiae*, including fission yeast, and in higher eukaryotes, Paf1C also contains an additional subunit, Ski8/Wdr61 (Figure 1) (Zhu et al., 2005a). Together with Ski3 and the RNA helicase Ski2, Ski8/Wdr61 is part of the conserved Ski complex, which functions as a general co-factor of the cytoplasmic exosome and assists with 3′-to-5′ RNA degradation (Halbach et al., 2013). It has been previously shown that in addition to the cytoplasm, the mammalian Ski complex localizes to actively transcribed genes in the nucleus in a Paf1C-dependent manner (Zhu et al., 2005a). Thus, Ski8/Wdr61 was suggested to link the two complexes, leading to speculation that mammalian Paf1C might orchestrate events downstream of RNA synthesis, such as RNA processing and quality control (Zhu et al., 2005a). However, so far these claims have not been substantiated by direct experimental evidence. Thus, **the first goal of my thesis was to revisit these observations using mouse embryonic stem cells (mESC) and elucidate the functional role of the interaction between the Paf1 and Ski complexes in mammalian nuclei.**

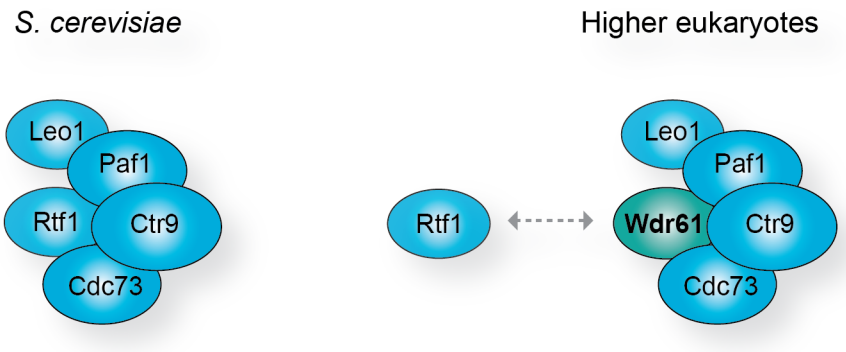



**Figure 1. Subunit composition of the Paf1 complex in *S. cerevisiae* and higher eukaryotes.** In *S. cerevisiae*, the Paf1 complex is composed of Ctr9, Paf1, Cdc73, Leo1 and Rtf1. In higher eukaryotes, Rtf1 is not stably associated with Paf1C and instead, the complex contains an additional subunit, Ski8/Wdr61, which is also part of the RNA helicase Ski complex functioning in cytoplasmic RNA decay.

## 2. The Paf1 complex is a repressor of RNAi-mediated transcriptional silencing

A novel role of Paf1C as a negative regulator of RNAi-mediated transcriptional repression was recently identified in a forward genetic screen in *S. pombe* (Kowalik et al., 2015). RNA interference (RNAi) is a gene silencing mechanism conserved in most eukaryotes and is broadly characterized by the ability of small double stranded RNAs to specifically repress the expression of complementary sequences (Meister and Tuschl, 2004). In the canonical RNAi pathway, Dicer processes double-stranded RNA into small interfering RNAs, which then guide Argonaute-containing effector complexes to specific targets by complementary base pairing (Meister and Tuschl, 2004). In mammalian somatic cells this process triggers post-transcriptional gene silencing through degradation of target mRNAs in the cytoplasm, without interfering with transcription. However, in organisms such as plants, *C. elegans*, ciliates and fungi, small RNAs also mediate transcriptional silencing of their targets by inducing DNA methylation or chromatin modifications in the nucleus (Grewal, 2010; Holoch and Moazed, 2015). This mechanism of small RNA-directed chromatin silencing is best understood in the fission yeast *S. pombe*, where constitutive heterochromatin at the pericentromeric regions of chromosomes is established and maintained by the RNAi machinery (Holoch and Moazed, 2015). In this organism endogenous small RNAs originating from centromeric repeat sequences are loaded into the RNA-induced Transcriptional Silencing complex (RITS), containing an argonaute protein (Verdel et al., 2004). RITS is then targeted to centromeric repeats via complementary base pairing between the Ago-bound siRNAs and the nascent transcript, leading to recruitment of the Clr4 histone methyltransferase (Bayne et al., 2010). Clr4 catalyzes methylation of histone H3 lysine 9

(H3K9me), which is a conserved feature of heterochromatin and serves as a binding site for the HP1 homolog Swi6 (Hayashi et al., 2012). Furthermore, target recognition by RITS recruits the RNA-dependent RNA polymerase (RdRP) which initiates double-stranded RNA synthesis using the nascent transcript target as a template. This double-stranded RNA substrate is further processed into small RNAs by Dicer thus leading to amplification of siRNAs (Colmenares et al., 2007; Motamedi et al., 2004). This process establishes a self-reinforcing positive-feedback loop between siRNA production and H3K9 methylation (Figure 2).

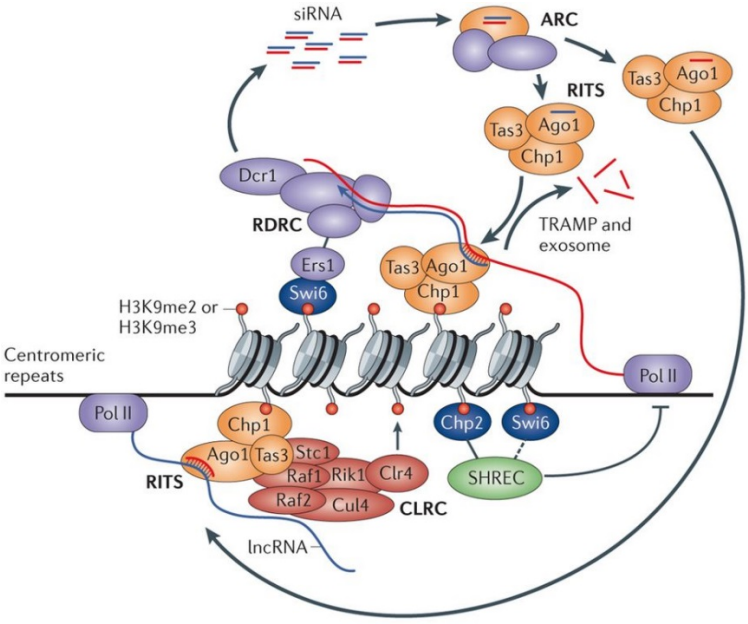



**Figure 2. Positive feedback loop between siRNA production and H3K9 methylation *in S. pombe.***
In *S. pombe*, endogenous siRNAs from centromeric repeat sequences are loaded onto the Ago-containing RITS complex (RNA-induced Transcriptional Silencing complex). RITS targets centromeric repeats via complementary base-pairing with the nascent transcript, leading to recruitment of the CLRC complex, which catalyzes methylation of histone H3 on lysine 9 (H3K9me), through the methyltransferase Clr4. Upon target recognition, RITS also recruits the RDRC complex, containing an RNA-dependent RNA polymerase (RdRP), which synthesizes double-stranded RNA using the nascent transcript as a template. These double-stranded RNA molecules are then processed by Dicer, leading to new production of siRNAs and further targeting of the RITS complex. Adapted with permission from (Holoch and Moazed, 2015).

A controversial topic in the RNAi field is whether the ability of siRNAs to trigger epigenetic silencing is conserved in mammalian somatic cells. Although there are a few studies that describe transcriptional

silencing in human cells (Kim et al., 2006; Morris et al., 2004; Suzuki et al., 2005), convincing evidence for a direct role of the RNAi machinery in specifying stable chromatin modifications is missing. Interestingly, *de novo* heterochromatin formation using synthetic siRNAs has been difficult to achieve even in organisms where the RNAi pathway has a well-established role in epigenetic silencing (Iida et al., 2008; Simmer et al., 2010). Our lab recently identified the Paf1 complex as a negative regulator of this process and showed that in Paf1C mutant strains, a synthetic hairpin RNA targeting a euchromatic gene is capable of triggering stable and heritable heterochromatin formation at the corresponding locus even in the absence of the initial siRNA source (Kowalik et al., 2015). Proposed models suggest that Paf1C represses siRNA-mediated heterochromatin formation by promoting efficient transcription termination and release of the nascent transcript. Altered transcription kinetics in Paf1C mutants leading to longer transcript retention on chromatin could potentially allow the RNAi machinery to establish more stable interactions with the target locus (Kowalik et al., 2015). Another model posits that Paf1C components oppose stable accumulation of the heterochromatic H3K9me mark by promoting rapid exchange of histone H3 (Sadeghi et al., 2015).

Given the high conservation of Paf1C from yeast to humans, it is possible that the complex performs similar repressive functions in higher organisms. This would provide novel opportunities to investigate whether RNAi-mediated epigenetic silencing is conserved in mammalian somatic cells. Therefore, **the second goal of my PhD project was to test if in the absence of the Paf1 complex, siRNAs can initiate *de novo* heterochromatin formation in mammalian cells, using mouse embryonic stem cells (mESCs) as a model system.**

# RESULTS

## 1. Generating Paf1C conditional knock out cells expressing endogenous siRNAs

To probe for a possible link between the Paf1 complex and RNAi-mediated chromatin modifications in mammalian cells, I used mouse embryonic stem cells (mESCs) as a model system. I took advantage of a previously generated *Leo1*$^{fl/fl}$ conditional knock out cell line expressing a CreERT2 recombinase fusion from the Rosa26 locus, where *Leo1* could be depleted upon addition of 4-hydroxytamoxifen (4OHT) to the culture media. I focused on this particular Paf1C component as it is not essential for viability in mESCs, allowing us to minimize pleotropic effects that might arise from depletion of essential subunits. Furthermore, Leo1 has a well establish role as a repressor of siRNA-mediated epigenetic silencing in *S. pombe* (Kowalik et al., 2015), making it a good candidate to manipulate in order to test a possible conservation of this process in mammalian cells.

My initial goal was to test if small RNAs originating from endogenous loci can direct heterochromatin formation at those regions (*in cis*) in the absence of a functional Paf1C. However, next generation sequencing experiments demonstrated that putative endogenous siRNAs (endo-siRNAs) are present at negligible amounts in mammalian somatic cells (Svoboda, 2014). Consistent with this, human Dicer cannot efficiently process long double-stranded RNA substrates, due to an auto-inhibitory function of its N-terminal helicase domains (Ma et al., 2008). Therefore, to induce endo-siRNA production, I integrated a mouse oocyte-specific Dicer isoform (Dicer$^{O}$) in the *Leo1*$^{fl/fl}$ conditional knock out cells. In contrast to somatic Dicer, Dicer$^{O}$ lacks the N-terminal DExD helicase domain (Figure 3A) and is able to efficiently process long duplex RNA substrates into small RNAs. Consequently, ectopic expression of Dicer$^{O}$ in mESC leads to the production of endo-siRNAs mapping to several unique genomic loci (Flemr et al., 2013). Using a random stable integration strategy, I generated two *Leo1*$^{fl/fl}$ Dicer$^{O}$ clones, expressing Dicer$^{O}$ at high and low levels (Figure 3B). Small RNA sequencing in these two clones confirmed the presence of endo-siRNAs, whose levels correlated with the amount of Dicer$^{O}$ expression (Figure 3C). I focused on endo-siRNAs originating from three unique genomic loci (*Anks3*, *Optn* and *Rad23a/Gadd45gip*), previously identified by Flemr and colleagues (Flemr et al., 2013) (Figure 3C). Small RNAs mapping to the *Anks3* gene are derived from long intramolecular interactions forming an RNA stem in intron 4. Similarly, endo-siRNAs match the 3′ UTR of *Optn*, which is predicted to fold into a duplex RNA structure. At the *Rad23a/Gadd45gip* locus, siRNAs map to a site of convergent transcription defined by the overlap between the 3′ UTRs of the two genes (Flemr et al., 2013). All three endo-siRNA clusters displayed the characteristic length and bias for the U-base at the starting position (Figure 3C), suggesting they represent *bona fide* siRNAs that can be loaded onto Ago proteins.

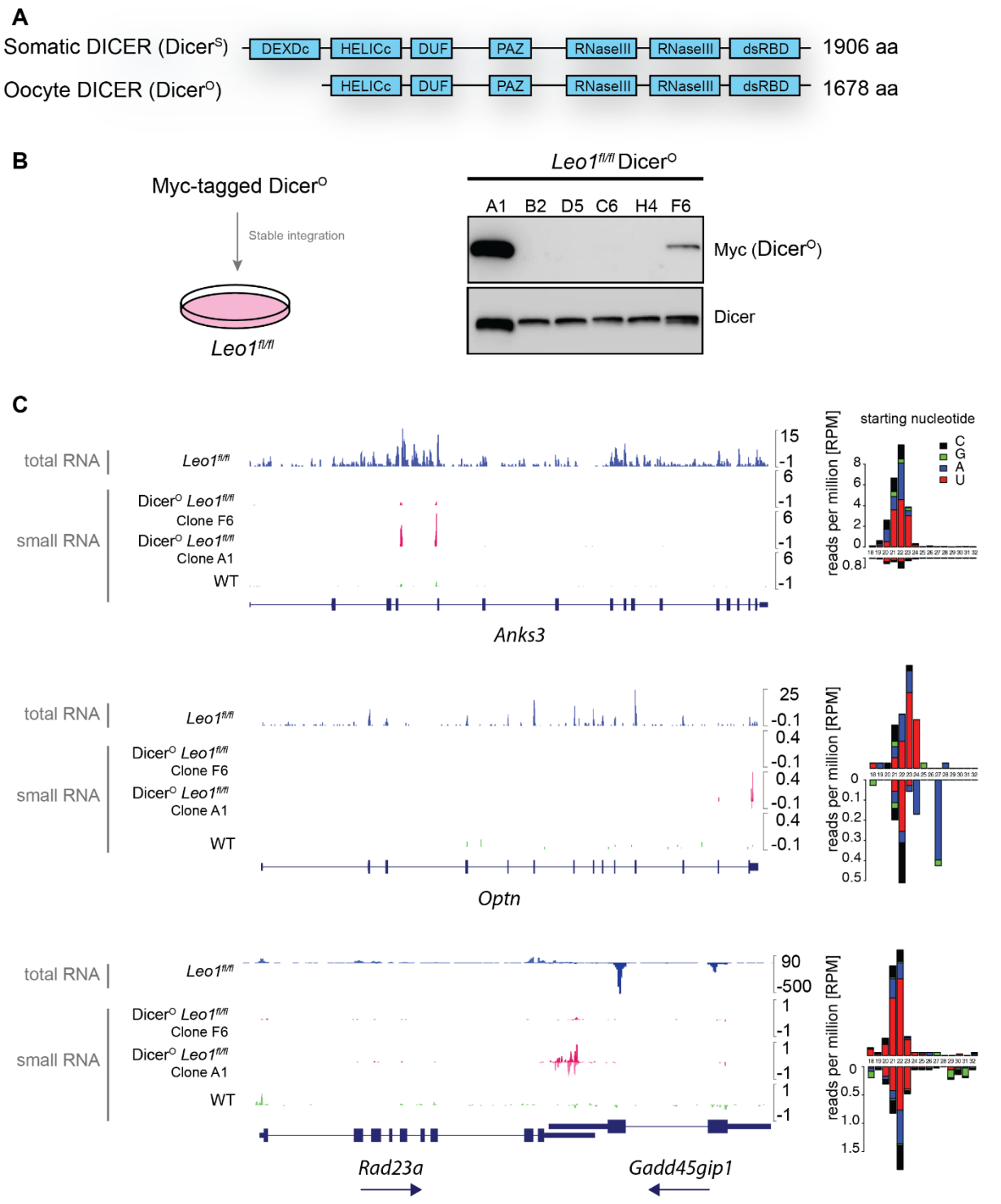



**Figure 3. Generating Paf1C conditional knock out cells expressing endo-siRNAs.**
**A** Schematic comparing the domain organization of somatic (Dicer$^S$) and oocyte Dicer (Dicer$^O$). **B** Left, outline of the strategy used to generate *Leo1*$^{fl/fl}$ cells with a stable integration of Dicer$^O$. Right, western blot analysis of Dicer$^S$ and Dicer$^O$ expression in the selected *Leo1*$^{fl/fl}$ Dicer$^O$ clones. **C** Left, total RNA-seq tracks of WT cells and small RNA-seq tracks (rpm) of three endo-siRNA-producing loci in *Leo1*$^{fl/fl}$ Dicer$^O$ clones A1 and F6. Right, length histograms of each endo-siRNA cluster. Clone A1 was used for all subsequent experiments.

## 2. Loss of *Leo1* affects the mRNA levels and histone modifications of the endo-siRNA-producing loci

After establishing mES cell lines that express endogenous siRNAs, I next asked whether loss of *Leo1* affects the expression and chromatin state of the three endo-siRNA-producing loci described above (Figure 3C). To induce *Leo1* knock out, I treated the *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells (clone A1, Figure 3) with 4OHT for up to eight days. *Leo1* mRNA levels were drastically reduced at day 2 and remained low throughout the time-course, confirming that *Leo1* was efficiently downregulated (Figure 4A). In the *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells, loss of *Leo1* led to 40-60% reduction of the mRNA levels of *Optn*, *Rad23a* and *Gadd45gip*, but not *Anks3* (Figure 4A). These observations would be consistent with three different scenarios. First, since Paf1C is a PolII elongation factor, its function might be required for efficient PolII transcription of the *Optn* and *Rad23a*/*Gadd45gip* loci. Second, endo-siRNAs derived from *Optn* and *Rad23a*/*Gadd45gip* map to the 3′UTR of these transcripts, thus potentially enabling post-transcriptional RNAi silencing in the cytoplasm. Third, if endo-siRNAs could indeed direct histone modifications, reduced mRNAs levels could potentially reflect formation of repressive chromatin environment at the target loci, leading to reduced transcriptional activity. To distinguish between these three scenarios, I next examined the mRNA expression levels of the above transcripts in *Leo1*$^{fl/fl}$ WT *Dicer* cells, where endo-siRNAs are not detectable by small-RNA sequencing. Similar to *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells, loss of *Leo1* in the WT *Dicer* background led to reduction of the mRNAs levels of *Optn*, *Rad23a* and *Gadd45gip* albeit to a lesser extent (Figure 4B), suggesting that the decrease in mRNA expression cannot be fully explained by the presence of Dicer$^{O}$ and endo-siRNAs.

To further explore if loss of *Leo1* affects the chromatin state of the endo-siRNA-producing loci, I treated the *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells with 4OHT for up to eight days to induce *Leo1* knock out and performed H3K9me2 ChIP-qPCR for each day of the 4OHT time course (Figure 4C). Methylation of histone H3 on lysine 9 is a hallmark of heterochromatin (Allis and Jenuwein, 2016) and in fission yeast, deletion of Paf1C subunits allows small RNA-mediated deposition of this repressive mark at homologous genes (Kowalik et al., 2015). Therefore, I speculated that in mammalian somatic cells, lack of *Leo1* could similarly enable accumulation of H3K9me2 at target loci in the presence of endo-siRNAs. ChIP-qPCR following a 4OHT time-course in the *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells showed that H3K9me2 levels at the *Anks3* and *Optn* loci (but not *Rad23a*/*Gadd45gip*) tend to increase throughout the time-course (Figure 4C). However, the same trend was observed after knocking out *Leo1* in a *Leo1*$^{fl/fl}$ WT *Dicer* cell line (Figure 4D), suggesting that H3K9me2 levels at the *Anks3* and *Optn* genes increase independently of endo-siRNAs.

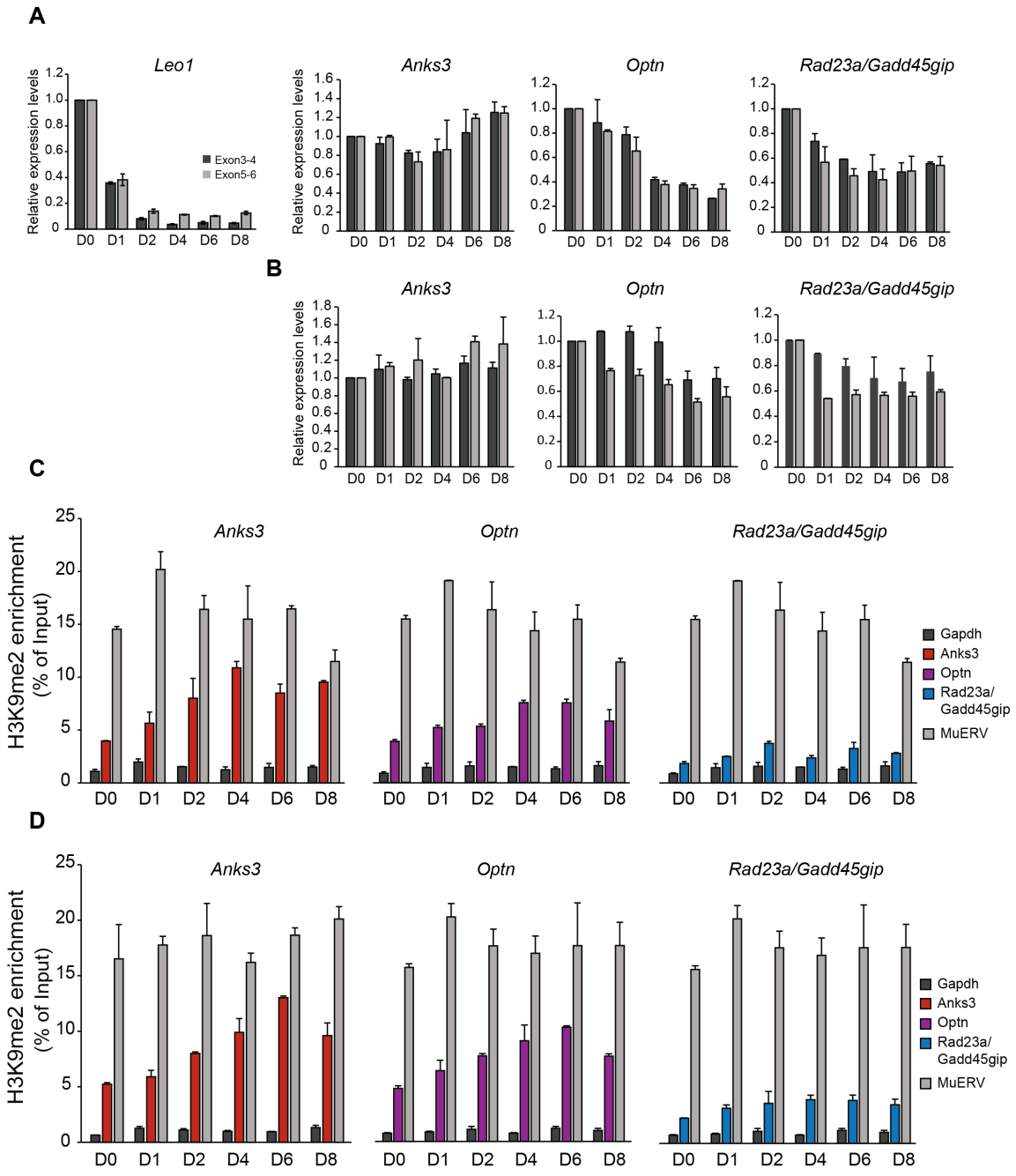



**Figure 4. Loss of Leo1 affects the mRNA levels and histone modifications of the endo-siRNA-producing loci.** **A** and **B** RT-qPCR measuring the mRNA expression levels of *Leo1*, *Anks3*, *Optn*, *Radd23a* and *Gadd45gip* in *Leo1*$^{fl/fl}$ Dicer$^{O}$ clone A1 (A) and *Leo1*$^{fl/fl}$ WT *Dicer* (B) cells throughout a 4OHT time-course. X-axes indicate the duration of the 4OHT treatment in days. Y-axes indicate relative mRNA levels normalized to TBP. Two different primer pairs were used for each gene, except for *Radd23a* and *Gadd45gip*, where black and grey bars correspond to the mRNA levels of *Radd23a* and *Gadd45gip*, respectively (See also Materials and Methods). **C** and **D** ChIP-qPCR measuring H3K9me2 levels at the *Anks3*, *Optn* and *Radd23/Gadd45gip* loci in *Leo1*$^{fl/fl}$ Dicer$^{O}$ clone A1 (C) and *Leo1*$^{fl/fl}$ WT *Dicer* (D) cells throughout a 4OHT time-course. X-axes indicate the duration of the 4OHT treatment in days. *Gapdh* and *MuERV* loci were used as negative and positive controls for H3K9me2 enrichment, respectively.

## 3. WT *Dicer* is not required for deposition of repressive chromatin marks at the Anks3 endo-siRNA producing region.

To exclude the possibility that WT *Dicer* produces very small amounts of endo-siRNAs, undetectable by sequencing but sufficient to induce the observed H3K9me2 changes upon *Leo1* depletion, I next used previously generated *Leo1*$^{fl/fl}$ cells where *Dicer* was knocked out (*Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$) and re-introduced WT *Dicer* in that background by stable integration. I treated the *Dicer* knock out and *Dicer* rescue cell lines with 4OHT for six days to deplete *Leo1* and performed ChIP-qPCR to measure the H3K9me2 levels at the *Anks3* locus. However, in contrast to previous observations, loss of *Leo1* did not correlate with accumulation of the H3K9me2 mark in the WT *Dicer* background (Figure 5B and Figure 4D). Similarly, H3K9me2 levels were not affected by the depletion of *Leo1* in either *Dicer*$^{-/-}$ or the WT *Dicer* rescue background (Figure 5B). Although not fully conclusive, these results argue against a possible direct link between endo-siRNAs and heterochromatin formation in a WT *Dicer* background.

## 4. Reporter silencing in the absence of *Leo1* is independent of siRNAs

The results obtained so far do not support a role for the mammalian Paf1C as an antagonist of RNAi-mediated heterochromatin formation in mESCs. However, in my experimental set up, the effects of disrupting Paf1C function are tested on the population level rather than on individual cells. Importantly, in *S. pombe* Paf1C mutants, silencing of siRNA-targeted euchromatic loci is efficiently maintained through mitotic divisions even in the absence of the original siRNA source (Kowalik et al., 2015). However, the rate at which individual yeast mutant cells initiate *de novo* heterochromatin formation and gene silencing is low (Flury et al., 2017). Therefore, it was possible that if similar to fission yeast, only a small subset of the *Leo1* knock-out mES cells could initiate heterochromatin formation, the effect will be largely masked at the population level. To address this, I engineered a fluorescent reporter, containing Neon Green fused to the *Anks3* siRNA-producing intron 4, driven by the endogenous *Anks3* promoter (Figure 6A). I chose *Anks3* since endo-siRNAs from this gene are not complementary to the mature transcript and would not trigger post-transcriptional silencing of the reporter mRNA in the cytoplasm. I then used TALENs to integrate a single copy of this reporter construct into a defined genomic site at the β-globin locus in both *Leo1*$^{fl/fl}$ Dicer$^{O}$ and *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ cells (Figure 6A). I isolated seven *Leo1*$^{fl/fl}$ Dicer$^{O}$ clones containing a homozygous integration of the Anks3 reporter, resulting in homogenous Neon Green expression as measured by FACS (Figure 6B). To assess how loss of *Leo1* affects expression of the reporter in individual *Leo1*$^{fl/fl}$ Dicer$^{O}$ cells, I performed a 4OHT time-course to induce *Leo1* knock out in all seven clones and monitored Neon Green levels by FACS. Interestingly, the proportion of non-fluorescent cells in all clones

gradually increased throughout the 4OHT time-course, although it remained below 1% (Figure 6C, left). I also noticed an inverse correlation between the basal Neon Green levels and the amount of non-fluorescent cells in each clone at the end of the 4OHT treatment (Figure 6C, right). This suggests that lower expression levels render the cells more susceptible to reporter silencing in the absence of a functional Paf1C.

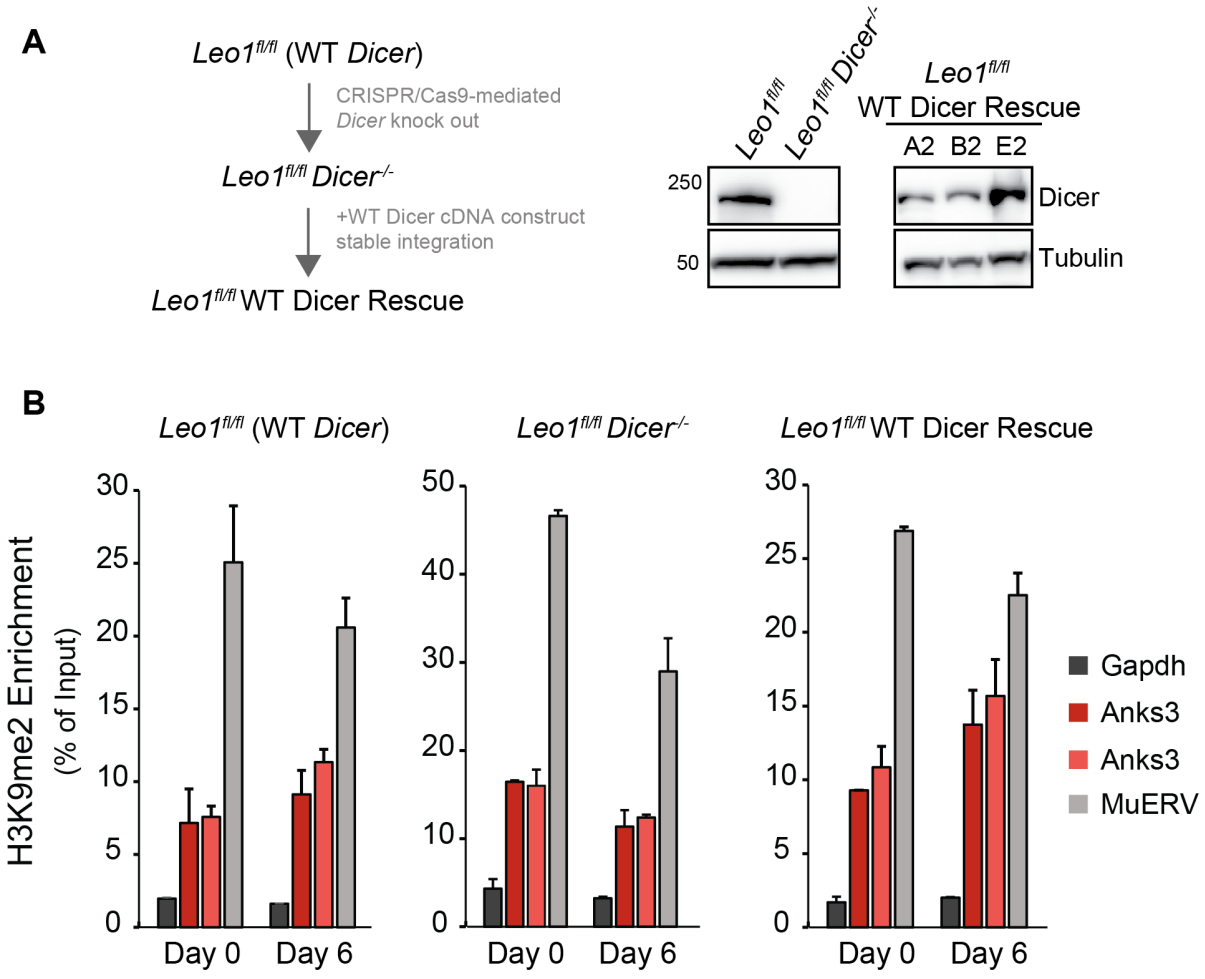



**Figure 5. WT Dicer is not required for deposition of repressive chromatin marks at the Anks3 endo-siRNA producing region.**

**A** Left, outline of the strategies used to generate *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ and *Leo1*$^{fl/fl}$ WT Dicer rescue cell lines. Right, western blot analysis of WT DICER protein levels in the indicated cell lines. *Leo1*$^{fl/fl}$ WT Dicer rescue clone A2 was used for subsequent ChIP experiments described in (B). **B** ChIP-qPCR measuring H3K9me2 levels at the *Anks3* endo-siRNA producing region in *Leo1*$^{fl/fl}$ (WT *Dicer*), *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ (*Dicer* knock out) and *Leo1*$^{fl/fl}$*Dicer*$^{-/-}$ WT Dicer rescue cell lines. X-axes indicate the duration of the 4OHT treatment in days. Dark- and light-red bars correspond to two different primer pairs that were used to measure H3K9me2 levels at the *Anks3* endo-siRNA producing region. *Gapdh* and *MuERV* loci were used as negative and positive controls for H3K9me2 enrichment, respectively.

To test whether the silent reporter state can be maintained and propagated in the absence of *Leo1*, I focused on *Leo1*$^{fl/fl}$ Dicer$^{O}$ reporter lines 2E, 3B and 7E (Figure 6C), as representatives of clones that exhibit low and high proportions of non-fluorescent cells upon depletion of *Leo1*. I isolated and expanded silent cells from the above clones which were either untreated (*Leo1*$^{+/+}$) or treated with 4OHT for 6 days (*Leo1* knock out) (Figure 6D). I then used this expanded population to assess the reporter expression status and chromatin state (Figure 6D). Interestingly, 4OHT-treated reporter lines 3B and 7E contained a much larger proportion of silent cells compared to their non-treated controls (Figure 6E). Western blot analysis of these cells confirmed that LEO1 was not expressed in the 4OHT-treated population (Figure 6F). However, compared to the parental cells, Dicer$^{O}$ protein levels were drastically reduced or almost completely undetectable in the reporter lines, irrespective of 4OHT (Figure 6F), suggesting these cells would not produce (many) siRNAs. Thus, although I observed a higher proportion of silent cells in the *Leo1* knock out reporter lines (Figure 6E), it is unclear whether this is due to siRNA-mediated transcriptional silencing.

To test whether the increased frequency of non-fluorescent cells in clones 2E, 3B and 7E (Figure 6E) correlates with higher levels of heterochromatin marks, I performed both H3K9me2 and H3K9me3 ChIP-qPCR using primers specific for the reporter locus (Figure 6G). Surprisingly, irrespective of 4OHT treatment, the $C_t$ values of input samples containing more silent cells were 2-7 cycles higher compared to those with lower proportions of silencing (Figure 6G). Since qPCR $C_t$ values should be very similar across all input samples, the above results could indicate the presence of micro-deletions in the reporter locus, that were selected for and propagated after FACS-sorting and expanding silent cells (Figure 6D). Therefore, in this case, downregulation of the fluorescent reporter could reflect impaired transcription, processing or translation.

My previous results indicate that WT *Dicer* likely does not have a direct effect on the chromatin state of the endo-siRNA-producing loci (Figure 5). Nevertheless, I wanted to revisit these observations using my reporter system. For this, I generated five *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ clones with a stable integration of the Anks3 reporter, resulting in homogenous Neon Green expression as assessed by FACS (Figure 6H). I then treated five *Leo1*$^{fl/fl}$ Dicer$^{O}$ and five *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ reporter lines with 4OHT for up to 6 days to deplete *Leo1*, and monitored Neon Green expression levels by FACS (Figure 6I, left). In line with my previous observations, the proportion of non-fluorescent cells in almost all *Leo1*$^{fl/fl}$ Dicer$^{O}$ clones increased throughout the 4OHT time-course. However, the number of silent cells in all five *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ reporter lines remained unchanged (Figure 6I, left). Although these observations are intriguing, I again noticed an inverse correlation between the initial Neon Green levels and the amount of non-fluorescent cells in each clone at the end of the 4OHT treatment. Importantly, all five *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ reporter lines had higher basal fluorescence than the *Leo1*$^{fl/fl}$ Dicer$^{O}$ clones with the largest proportion of silent cells (Figure 6I, left). This

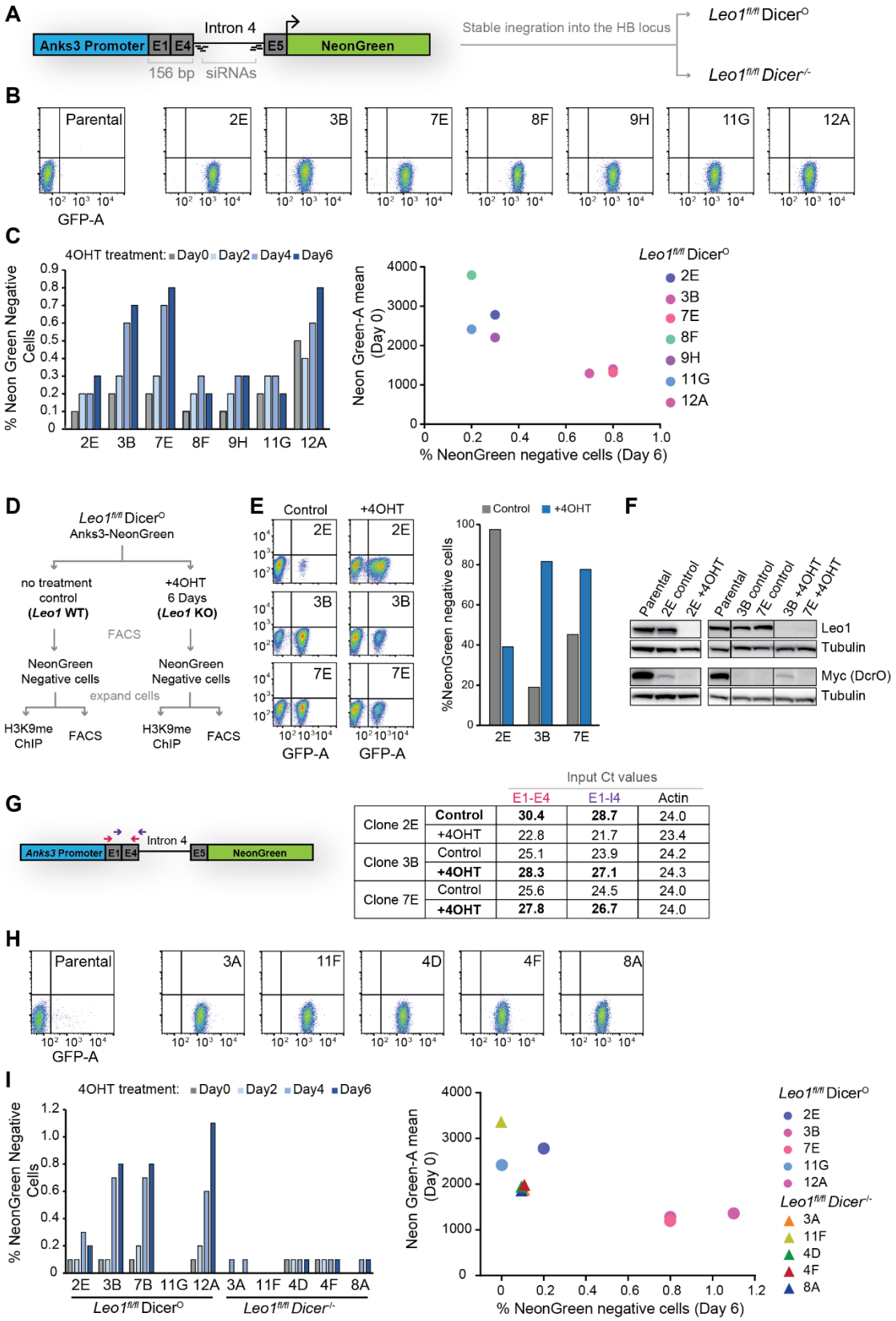



**Figure 6. Reporter silencing in the absence of *Leo1* is independent of siRNAs.**

**Figure 6. Reporter silencing in the absence of *Leo1* is independent of siRNAs *(continued)***
**A** Illustration of the Anks3 Neon Green reporter construct. **B** Flow cytometry analysis of Neon Green expression in seven *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter clones with a stable integration of the reporter described in (A). The parental *Leo1*$^{fl/fl}$ Dicer$^{O}$ cell line from which all seven clones are derived is shown as an example of non-fluorescent cells. **C** Left, percent Neon Green negative cells determined by flow cytometry analysis of the seven *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter clones from (B) during a 6-day 4OHT time-course treatment. Right, average fluorescence intensity vs. percent Neon Green negative cells at 4OHT day 6 for the indicated *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter lines. **D** Outline of the experimental set up used to test the maintenance of reporter silencing upon deletion of *Leo1*. **E** Left, Flow cytometry analysis of Neon Green expression in the indicated *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter lines derived from previously FACS-sorted and expanded silent cells. Right, quantification of the percent Neon Green negative cells from the FACS profiles of the indicated clones. **F** Western blot analysis of LEO1 and myc-tagged DICER$^{O}$ protein levels in the *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter lines from (E). **G** Left, ChIP-qPCR primer binding sites along the Anks3 Neon Green reporter. Right, $C_t$ values of input samples from H3K9me2 and H3K9me2 ChIP-qPCR of the *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3 reporter lines from (E). **H** Flow cytometry analysis of Neon Green expression in five *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3 reporter clones with a stable integration of the reporter described in (A). A parental *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ cell line is shown as an example of non-fluorescent cells. **I** Left, percent Neon Green negative cells determined by flow cytometry analysis of the indicated reporter cell lines during a 6-day 4OHT time-course treatment. Right, average fluorescence intensity vs. percent Neon Green negative cells at 4OHT day 6.

makes it difficult to directly compare the WT *Dicer* and Dicer$^{O}$ backgrounds, since variation in basal reporter expression could account for the observed differences in reporter silencing upon *Leo1* knock out. Collectively, these data indicate that initiation and maintenance of reporter silencing is likely not dependent on siRNAs.

## 5. Mammalian Paf1 complex does not interact with the Ski complex in mES cells

It has been previously suggested that in addition to its function in transcription, the human Paf1 complex might also orchestrate RNA processing and quality control, based on its reported interaction with the RNA helicase Ski complex (Zhu et al., 2005a). The Ski complex is primarily cytoplasmic and has a well-established role as a general co-factor of the cytoplasmic RNA exosome that assists with 3′-to-5′ RNA degradation (Halbach et al., 2013). However, a previous study suggested that in human cells, the Ski complex also localizes to the nucleus, where it interacts with Paf1C through the common subunit WDR61 (Zhu et al., 2005a). Since it is still unclear whether the two complexes function together, I sought to address this question using mES cells. I created cell lines with endogenously tagged 3xFLAG-AviTag WDR61 and SKIV2L in a background expressing a promiscuous biotin ligase (BirA ligase). Western blot analysis showed that the two tagged proteins were properly expressed and biotinylated (Figure 7A). I then used our endogenously tagged cells and performed streptavidin pull downs followed by mass spectrometry. Using the common WDR61 subunit as a bait, I identified all Paf1C components (LEO1, CTR9, PAF1 and CDC73) as well as the Ski complex subunits TTC37 and SKIV2L (Figure 7B, left). However, streptavidin

purification of the Ski complex subunit SKIV2L, revealed interactions with the rest of the Ski complex components TTC37 and WDR61, but not with any Paf1C subunits (Figure 7B, right). I next performed co-IP experiments as a complementary approach to IP-MS. Using a FLAG immunoprecipitation of endogenously tagged SKIV2L, I did not detect any interaction with Paf1C subunits CTR9, LEO1 or PAF1 (Figure 7C). Taken together these data indicate that the mammalian Ski complex and Paf1C do not form a stable interaction although they contain a common subunit.

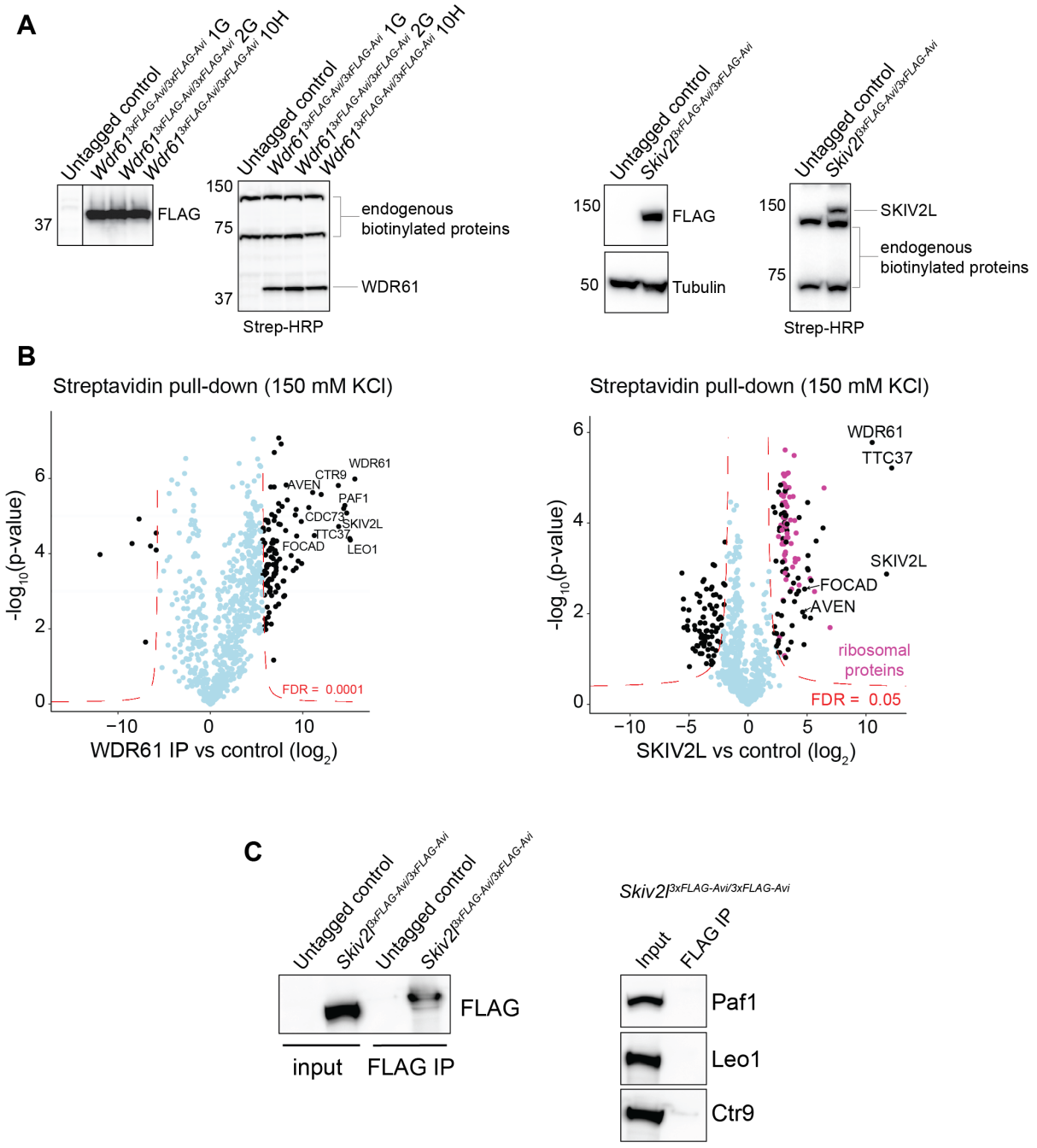



**Figure 7. The mammalian Paf1 complex does not interact with the Ski complex.**

**A** Western blots showing expression and biotinylation of endogenously tagged 3xFLAG-AviTag WDR61 and SKIV2L. **B** Mass spectrometry analysis of single step streptavidin purification of endogenously tagged WDR61 (clone 1G) and SKIV2L using 150mM KCl. All experiments included three technical replicates and an untagged mESC as a control. FDR, false discovery rate. **C** Western blot analysis of endogenously tagged 3xFLAG-AviTag SKIV2L co-IP experiments.

# DISCUSSION

Although initially identified as a posttranscriptional silencing mechanism (Fire et al., 1998), RNAi pathways also play important roles in epigenetic gene regulation in fungi, ciliates, plants and *C. elegans* (Grewal, 2010; Holoch and Moazed, 2015). Whether siRNAs can specify histone and DNA modifications in mammalian cells however, remains somewhat controversial. In fact, initiating silent histone modifications at euchromatic loci using synthetic siRNAs has been difficult to achieve even in organisms where RNAi has a well-established role in epigenetic phenomena (Buhler et al., 2006b; Iida et al., 2008; Simmer et al., 2010). The recent discovery of the Paf1 complex as a repressor of siRNA-mediated heterochromatin formation in *S. pombe* (Kowalik et al., 2015) has opened exciting new avenues to explore a possible conservation of this process in mammalian cells. To test this hypothesis, I asked whether endogenous siRNAs produced by ectopic expression of oocyte Dicer$^{O}$ in mES cells (Figure 3) can induce epigenetic silencing of their cognate loci (*in cis*) upon conditional ablation of the Paf1C subunit *Leo1*. Although the mRNA expression of some siRNA-producing loci decreased while the H3K9me2 levels showed a tendency to increase upon loss of *Leo1*, this trend was not correlated with the presence of Dicer$^{O}$ (Figure 4). These findings suggest that the observed changes were not mediated by RNAi since endo-siRNAs were not detected in cells expressing only WT *Dicer* (Figure 3). Furthermore, I ruled out the possibility that WT DICER generates low levels endo-siRNAs that could account for the accumulation of the H3K9me2 (Figure 5). I also note that H3K9me2 levels were measured by ChIP-qPCR specifically along the DNA sequences where endo-siRNAs map. Thus, regardless of the underlying trigger, it is possible that the observed increase in H3K9me2 levels upon *Leo1* knock out is more localized and thus insufficient to confer transcriptional repression. Furthermore, in the case of *Anks3* and *Optn*, H3K9me2 levels increased only up to two or three fold above basal levels (Figure 4), which is unlikely to result in robust transcriptional changes.

As a complementary approach to my experiments on the population level, I designed a fluorescent reporter based on the *Anks3* endo-siRNA producing sequence, allowing me to monitor initiation and maintenance of silencing in individual cells using FACS. This allowed me to assess whether similar to *S. pombe*, only a subset of mES cells were able initiate RNAi-mediated transcriptional silencing in the absence of a functional Paf1C. Furthermore, unlike the endogenous *Anks3* locus, the siRNA-generating intron was positioned in close proximity to the promoter of the reporter construct (Figure 5A). I reasoned that this feature might render the promoter more susceptible to silencing even if the endo-siRNAs were able to induce only limited and localized chromatin changes. I initially found that the proportion of non-fluorescent cells in the *Leo1*$^{fl/fl}$ Dicer$^{O}$ background increased in the absence of *Leo1* (Figure 6C). However, further

experiments testing the maintenance of silencing, revealed that Dicer$^{O}$ protein levels were significantly reduced or completely undetectable in cultured non-fluorescent cells (Figure 6E and F). This argues against a direct link between siRNAs and reporter silencing. Furthermore, ChIP-qPCR experiments hint at the presence of genomic abnormalities at the reporter locus of cell populations which maintain a higher frequency of silencing (Figure 6G). This suggests that reporter repression is more likely a consequence of abnormal transcription, processing and/or translation, rather than a direct result of RNAi-induced transcriptional silencing. I also note that although the reporter was expressed from a defined genomic locus, there were clonal variations in basal fluorescence levels (Figure 6C and Figure 6I), which made it difficult to directly compare silencing in different backgrounds (e.g. *Dicer*$^{-/-}$ vs. Dicer$^{O}$). Therefore, a reporter cell line with an inducible Dicer$^{O}$ and/or siRNA-source might be a more suitable approach to address this question.

Overall, I did not find convincing evidence that siRNAs can direct heterochromatin formation and transcriptional silencing in mES cells lacking a functional Paf1 complex. However, further experiments are necessary to definitively rule out the presence of such mechanism in mammalian cells. For instance, my experimental set up relies on small amounts of endogenous siRNAs (Figure 3C), which might not be sufficient to trigger robust chromatin changes leading to transcriptional silencing. To address this, it might be necessary to provide a more potent siRNA source such as a synthetic hairpin or other long double-stranded RNA substrates in combination with Dicer$^{O}$. This approach should be feasible, especially in stem cells, where attenuated interferon responses to double-stranded RNA (Chen et al., 2010b; Pare and Sullivan, 2014; Wang et al., 2013) should minimize the toxicity that is usually associated with such exogenous substrates. Similarly, it is possible that certain features of our Anks3 reporter construct might be incompatible with robust siRNA-mediated transcriptional silencing. For instance, I purposefully positioned the Anks3 endo-siRNA intron close to the reporter promoter. However, while recent mechanistic studies in *S. pombe* revealed that intronic siRNA targets are compatible with RNAi-directed heterochromatin formation, it was also shown that targeting promoter-proximal introns can compromise the initiation and maintenance of silencing (Shimada et al., 2016). Likewise, it is possible that my reporter construct was less susceptible to repression, due to the close proximity of the promoter and the endo-siRNA-producing intron (Figure 6). Furthermore, several studies describing RNAi-mediated epigenetic silencing in human cells have used siRNAs specifically targeting promotor sequences (Kim et al., 2006; Morris et al., 2004; Suzuki et al., 2005) and it would be important to test if these results can be reproduced in my experimental system using promoter siRNAs.

Since mammalian cells lack an RNA-dependent RNA polymerase (de Farias et al., 2017; Zong et al., 2009) which is required to amplify production of the siRNAs, it might be interesting to test whether ectopic

expression of an RdRP in combination with Dicer$^{O}$ and an siRNA source could induce transcriptional silencing in Paf1C deficient mES cells. However, if RNAi-mediated epigenetic silencing can be achieved only under such artificial conditions, it is highly unlikely that normal biological processes in mammalian somatic cells would also depend on this mechanism. Murine *Dicer* is essential for embryonic development (Bernstein et al., 2003) and it has been previously reported that centromeric repeat sequences are derepressed in *Dicer*-deficient mES cells due to defects in centromeric heterochromatin formation, leading to suggestions that the RNAi machinery might be directly involved in this process (Kanellopoulou et al., 2005). However, these results could reflect secondary effects caused by global changes in gene expression, especially since miRNA biogenesis is also perturbed in the absence of *Dicer*.

Conflicting evidence about the cellular localization of RNAi components has further fueled the controversy surrounding siRNA-mediated epigenetic silencing in mammalian cells. While some studies report that the RNAi factors are confined to the cytoplasm (Much et al., 2016; Stalder et al., 2013; Zeng and Cullen, 2002), others suggest that the RNAi machinery has both cytoplasmic and nuclear localization (Burger and Gullerova, 2018; Doyle et al., 2013; Ohrt et al., 2012; Robb et al., 2005; White et al., 2014). Another study reports that the nuclei of human somatic cells contain active RNAi factors, but no RISC loading components, suggesting that siRNAs cannot be loaded in the nucleus (Gagnon et al., 2014). Thus, questions about the presence and significance of a nuclear RNAi pathway in mammalian cells remain subject of debate. In the present study, I also found no supporting evidence that siRNA-mediated epigenetic silencing in mammalian cells can occur in the absence of Paf1C.

Differences in Paf1 complex composition in yeast versus higher eukaryotes prompted us to investigate whether mammalian Paf1C activity extends beyond transcription. In mammalian cells, Paf1C contains a novel subunit, Wdr61, which is also part of the conserved RNA helicase Ski complex, functioning in RNA decay (Zhu et al., 2005a). This has led to suggestions that in the nucleus, the two complexes interact through this common subunit, although so far no specific function has been ascribed to that interaction. To address this, I first sought to confirm that WDR61 links the Paf1 and Ski complexes in mES cells. However, in contrast to observations by Zhu et al., my IP-MS data clearly indicate that the two complexes do not interact stably. Although I cannot exclude a more transient association or discrepancy due to differences in cell lines (human vs. mouse), my results are consistent with the fact that Paf1C and the Ski complex perform their canonical functions in transcription and cytoplasmic RNA decay in different cellular compartments. Moreover, the yeast Ski complex was recently reported to bind directly to a translating 80S ribosome during RNA degradation (Schmidt et al., 2016), thus providing further support for the cytoplasmic localization of the Ski complex. Although I found no evidence that the two complexes can function together, further

investigation of the mammalian Ski complex revealed novel interaction partners and interesting insights into translation-coupled RNA decay, which are the focus of the second chapter of this thesis.

# MATERIALS AND METHODS

**Experimental model**

Male 129×C57BL/6 mouse embryonic stem cells (mESC) (Mohn et al., 2008) were grown in serum/LIF media (DMEM (Gibco 21969-035) supplemented with 15% fetal bovine serum (Gibco 10270106), 2 mM L-glutamine (Gibco 25030024), 1x non-essential amino acids (Gibco 11140035), 1 mM sodium pyruvate (Gibco 11360070), 0.1 mM 2-mercaptoethanol (Sigma M-7522), 50 mg ml−1 penicillin, 80 mg ml−1 streptomycin and homemade LIF) at 37 °C in 5% CO2. Cells were cultured on dishes coated with 0.1% gelatin (Sigma G1890).

**Generation of endogenously tagged cell lines**

Endogenous gene tagging with a 3xFLAG-AviTag was performed using TALEN homology-directed repair with single-stranded oligodeoxynucleotide (ssODN) donor templates encoding the tag, flanked by 5′ and 3′ homology arms. The ssODNs donors were synthetized as ultramers by Integrated DNA Technologies. N-terminally tagged *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 8F (cMB331) and *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 1G (AR90) were generated using TALENs cutting near the start codon. *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 8F (cMB331) was generated in a mES 129×C57BL/6 cells expressing BirA ligase and CreERT2 from the Rosa26 locus (cMB063) (Ostapcuk et al., 2018). *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 1G (AR90) was generated in mES 29×C57BL/6 cells expressing only BirA ligase from the Rosa26 locus (cMB053) (Ostapcuk et al., 2018). All clones were screened for homozygous integration of the tag by PCR and Sanger sequencing and expression of the fusion proteins was confirmed by western blot with a FLAG antibody. Biotinylation of the tag was verified by western blot using streptavidin-HRP. See also "List of cell lines" table at the end of the Materials and Methods section.

**Generation of cell lines with stable integrations**

*Leo1*$^{fl/fl}$ Dicer$^{O}$ clones A1 and F6 were generated by random integration of DicerO-MycHis cDNA construct (Flemr et al., 2013) in a *Leo1*$^{fl/fl}$ background expressing CreERT2 from the *Rosa26* locus (Clone 5C, cMB068). DicerO expression is driven by the human EF1α promoter. The cells were selected with 10μg/mL blasticidin for 5 days. Clones A1 and F6 were used for initial small RNA-seq experiment. Clone A1 was used for all subsequent experiments. *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ WT *Dicer* rescue cell lines were generated by random integration of Dicer-MycHis cDNA construct (Flemr et al., 2013) in a *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ background (cMB081) using the same random integration strategy as for Dicer$^{O}$. Only *Leo1*$^{fl/fl}$ Dicer$^{-/-}$ WT *Dicer* rescue clone A2 was used for subsequent ChIP experiments. To induce *Leo1* knock out, the cells were treated with 0.1 μM 4-hydroxytamoxifen (4-OHT) (Sigma) for 0, 1, 2, 4, 6, or 8 days (see Figure 4A for *Leo1* mRNA

expression levels during the time course). See also "List of cell lines" table at the end of the Materials and Methods section.

**Generation of Anks3-Neon-Green reporter cell lines**

The Anks3-Neon-Green reporter, driven by the endogenous mouse *Anks3* promoter was integrated into a defined site at the β-globin locus in both *Leo1*$^{fl/fl}$ Dicer$^{O}$ (Clone A1) and *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ (cMB079 and cMB081) cells. The reporter and its promoter were cloned into a pBlu-Neon-Green plasmid between the 5′ and 3′ β-globin locus homology arms. The resulting reporter plasmid was then used as donor for homology-directed repair of a dsDNA cut at the β-globin locus induced by TALENs. For specific primer and TALEN sequences and a list of cell lines, see the tables at the end of the Materials and Methods section.

**Transfections**

For endogenous tagging of *Skiv2l* (cMB331) and *Wdr61* (AR90) with TALENs, cells were transfected with 400 ng of each TALEN, 1000 ng ssODN donor and 100 ng of pRRP puromycin recombination reporter (Flemr and Buhler, 2015). 24 hours post-transfection, the cells were selected with 2µg/ml puromycin for 28 hours and surviving cells were plated at clonal densities as described above. For integration of the Anks3-Non-Green reporter into the β-globin locus, cells were transfected with TALEN 400 ng of each TALEN, 1000 ng on the pBlu-Anks3-Neon-Green plasmid donor and 100 ng of pRRP puromycin recombination reporter. For genome editing with TALENs, 24 hours post-transfection, the cells were selected with 2µg/ml puromycin for 28 hours and surviving cells were seeded sparsely on 10 cm plates for clonal expansion. After 5-7 days, colonies were individually picked into 96-well plates, expanded and genotyped by PCR. Cells with proper in-frame homozygous insertions of the 3xFLAG-Avi tag were further confirmed by Sanger sequencing and western blot and Anks3-Neon-Green reporter lines were analyzed by FACS for homogenous Neon Green expression. For WT *Dicer* and Dicer$^{O}$ random integrations, cells were transfected with 2µg of cDNA plasmids. 24-hours post-transfection, the cells were selected with 10µg/mL blasticidin for 24 hours and surviving cells were seeded sparsely on 10 cm plates for clonal expansion for 5-7 days in the presence of 10µg/mL blasticidin. Individual clones were then picked picked into 96-well plates, expanded and screened by western blot with an anti-Myc antibody for Dicer$^{O}$ integration and an anti-dicer for WT *Dicer* integration. All transfections were carried out with Lipofectamine 3000 reagent (Invitrogen L3000015) at 3 µl per 1 µg of total DNA in OptiMem media (Gibco, 31985062). Approximately 500,000 cells were used for each transfection.

**Cloning**

The Anks3-Neon-Green reporter contains the Promoter region of Anks3 (approximately 1000bp upstream of the annotated transcription start site) and portion of Exon1, fused to the siRNA-producing Intron 4, flanked by portions of Exon4 and Exon5. The Promoter-Exon fragment was amplified from mESC gDNA with a forward primer containing SpeI site and a reverse primer containing AatII site from, and the Exon4-Intron4-Exon5 sequence was amplified with a forward primer containing an AatII site and a reverse primer containing AscI site. The resulting fragments were cloned into a NheI/AscI-digested pBlu-HB-Neon-Green plasmid containing 5′ and 3′ homology arms for targeted integration into the β-globin locus (plasmid pMB2032). For specific primer sequences, see the "List of primers" table at the end of the Materials and Methods section.

**Small RNA sequencing**

Approximately 500,000 cells were seeded in a 6-well plate and cultured for 48 hours. The cells were trypsinized, collected in media, pelleted and washed with PBS. Small RNAs were isolated with mirPremier microRNA Isolation Kit from Sigma (SNC50). Libraries were prepared from 500ng small RNAs using the Illumina TruSeq Small RNA kit and sequenced on an Illumina HiSeq2500 platform (50 nt single-end reads).

**Total RNA sequencing**

Total RNA from a nearly confluent 6-well plate of *Leo1*$^{fl/fl}$ cells was isolated using the Agilent Absolutely RNA Miniprep Kit with on-column DNase digestion. After ribosomal RNA depletion with the Illumina Ribozero kit, libraries were constructed using ScriptSeq v2 (Epicentre; SSV21106) and sequenced on an Illumina HiSeq2500 platform (50 nt single-end reads).

**RT-qPCR**

Total RNA was extracted from mES cells with the Agilent Absolutely RNA Miniprep Kit and 500ng of RNA was reverse-transcribed using the PrimeScript RT-PCR Kit (Takara Bio RR036A-1). qPCR was performed with SsoAdvanced SYBR Green Supermix (Bio-Rad 172-5274) on a CFX96 Real-Time PCR System (Bio-Rad) and relative RNA levels were calculated using the ΔCt method and normalization to TBP mRNA abundance). See also "List of cell lines" table at the end of the Materials and Methods section.

**Chromatin immunoprecipitation (ChIP)**

H3K9me2 chromatin immunoprecipitations described in Figure 4 and Figure 5 were carried out as follows: One 10cm dish of confluent cells per sample was fixed in media supplemented with 1% Formaldehyde for 10 min at room temperature, quenched with glycine for 8 min at room temperature and washed with PBS.

The cells were harvested in PBS supplemented with 0.5% BSA and 1X protease inhibitor cocktail (Roche) by scraping and cell pellets were flash-frozen in liquid nitrogen and stored at -80 °C. To prepare nuclear fractions, the cells were resuspended in 8mL buffer NP-Rinse1 (50mM Hepes, pH 8.0, 150mM NaCl, 1mM EDTA, 10% Glycerol, 0.5% NP-40, 0.25% Triton-X100) and incubated on a rotating wheel at 4 °C. The cells were spun down for 5 min at 1,2000g at 4 °C and the pellet was resuspended in in 10mL buffer NP-Rinse2 (10mM Tris, pH 8.0, 1mM EDTA, 200mM NaCl) and spun down at 1,2000g for 5 min at 4 °C. The supernatant was removed, and the pellet was washed two times with 5mL shearing buffer (10mM Tris, pH 8.0, 1mM EDTA, 0.1% SDS) and resuspended in 900 µl shearing buffer, supplemented with 1X protease inhibitor cocktail (Roche). The samples were then transferred to 12x12 AFA Covaris tubes, topped with shearing buffer to the rim and sonicated in Covaris S220 with a water bath filled with 1.52L of dd$H_2O$, using the following sonication conditions: 5% Duty cycle, 200 cycles per burst, continuous degassing mode, 140W peak incident power, 4 °C bath temperature.

After sonication, the samples were transferred to 1.5mL tubes and clarified by centrigugation at 16,000 xg for 10 min at 4 °C. Protein concentrations was measured using the BioRad protein assay and for each sample approximately 500µg of total protein in 500µl was added to 125µl 5X IP buffer (250mM Hepes, pH 7.5, 750mM NaCl, 1mM EDTA, 0.1% SDS, 0.5% Na-Deoxycholate, 5% triton) for a final concentration of 1X IP buffer (50mM Hepes, pH 7.5, 150mM NaCl, 1mM EDTA, 0.1% SDS, 0.1% Na-Deoxycholate, 1% triton). The samples were then incubated with 2µg of H3K9me2 antibody (Active Motif) overnight at 4 °C. Before addition of the antibody, approximately 5% of the sample volume was set aside as an input sample. On the following day, the IP samples were clarified by centrifugation at 16,000g for 5 min and transferred to 25µl M-280 Sheep Anti-Mouse IgG Dynabeads (Invitrogen) prewashed with 1X IP buffer, and incubated for 3 hours at 4 °C on a rotating wheel. The beads were then washed twice with 1X IP buffer, once with DoC wash buffer (10mM Tris, pH 8.0, 1mM EDTA, 0.25M LiCl, 0.5% NP-40, 0.5% Na-Deoxycholates) and once with 1X TE. Samples were eluted twice by resuspending the beads in 50µl elution buffer (1% SDS, 0.1M $NaHCO_3$), following an incubation at 65°C for 20 min at 1000 rpm. Input samples were adjusted to the same volume (100µl) with elution buffer and processed together with the IPs. All samples were incubated with 2µg RNase A and incubated for 30 min at 37°C. After that the samples were treated with 20µg Proteinase K for 2.5 hours at 55°C and de-crosslinked for 7 hours at 65°C. DNA was isolated by phenol/chlorophorm, precipitated with ethanol and dissolved in 50µl $H_2O$. The purified DNA was then subjected to qPCR analysis using ChIP primers.

H3K9me2 and H3K9me3 chromatin immunoprecipitations described in Figure 6 was carried out without crosslinking as follows: Approximately $1x10^6$ cells for each sample were trypsinized, collected in media

and washed once with cold PBS. To isolate the nuclei, the cells were resuspended in 1ml ice cold PBS containing 0.1% NP-40 and incubated 5 min on ice. The nuclei were briefly spun down and washed once with 500µl cold MNase buffer (10mM Tris-HCl, pH 7.5, 4mM $MgCl_2$, 1mM $CaCl_2$) and resuspended in in 250µl cold MNase buffer supplemented with 1X protease inhibitor complex (Roche). The chromatin was digested with 0.25µl MNase (NEB, 2000U/ul) for 7 min at 37ºC, 1200rpm. The reaction was stopped by adding 25ul of STOP solution (166mM EDTA, 333mM EGTA) and incubating the samples on ice for 5 min. The nuclei were lysed in 1ml Hypotonic lysis buffer (0.2mM EDTA, 1.5mM DTT, 20mM EGTA) , supplemented with 1X protease inhibitor cocktail (Roche) for 1 hour on ice with occasional vortexing. Insoluble debris were pelleted by centrifugation at 13,000g for 5 min at 4ºC and 125µl of 10X N-ChIP buffer (100mM Tris-HCl, pH 8.0, 20mM EDTA pH8.0, 200mM NaCl, 10% Triton-X100, 1% SDS, 10X protease inhibitor cocktail, Roche). The samples were split in half and incubated with either 3µg of mouse anti-H3K9me2 antibody (Wako, 302-32369) or 3µg of rabbit anti-H3K9me2 (Abcam ab8889) overnight at 4ºC on a rotating wheel. Prior to addition of the antibodies, 5% of each sample was set aside as an input. On the following day, H3K9me2 IPs were incubated with 40µl of Protein G Dynabeads (Invitrogen) and H3K9me3 IPs were incubated with a 1:1 mix of Protein A or G Dynabeads (Invitrogen) for 2 hours at 4ºC on a rotating wheel. The beads were washed twice in RIPA buffer (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 140mM NaCl, 1% Triton-X100, 0.1% SDS, 0.1% Sodium Deoxycholate), twice in RIPA500 buffer (10mM Tris-HCl pH 8.0, 1mM EDTA pH 8.0, 500mM NaCl, 1% Triton-X100, 0.1% SDS, 0.1% Sodium Deoxycholate), twice with LiCl, 1X TE. DNA was eluted twice in 50µl Elution buffer (10mM Tris-HCl pH 8.0, 5mM EDTA pH 8.0, 150mM NaCl, 1% SDS) supplemented with 20µg Proteinase K, for 30 min at 55ºC, 1400rpm. Input samples were adjusted to 150µl with elution buffer and 20µg Proteinase K, and incubated for 1 hour at 55ºC, 1400rpm (in parallel with elution). DNA was purified with 150ul phenol/chloroform mix, precipitated with ethanol and resuspended in 50 µl 10mM Tris pH8.5. The purified DNA was then subjected to qPCR analysis using ChIP primers). See also "List of cell lines" table at the end of the Materials and Methods section.

**Flow cytometry**

Cells grown on 6-well plates were harvested by trypsinization, collected in media and washed two times with PBS. Neon Green expression was analyzed (20,000 cells per sample) on BD LSRII SORP Analyser (Becton Dickinson). For experiments testing the maintenance of reporter silencing, Neon-green negative cells harvested from 10cm dishes as described above were sorted on BD FACSAria cell sorter (Becton Dickinson). After expanding the sorted Neon-Green negative population, the Neon-Green expression levels were analyzed on BD LSRII SORP Analyser (Becton Dickinson) as described above.

**Western Blotting**

Cells were lysed for 30 min on ice in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 % Triton-X, 0.5 mM EDTA, 5 % glycerol, 1x protease inhibitor cocktail (Roche) and 1 mM DTT. Lysates were clarified by centrifugation at 16,000 xg for 10 min at 4 °C and protein concentration was measured using the BioRad protein assay. Approximately 20 µg of total protein extract was resolved on NuPAGE-Novex Bis-Tris 4–12 % gradient gels (Thermo Fisher), transferred semi-dry to a polyvinylidene fluoride (PVDF) membrane, blocked in 5 % non-fat milk in TBS+0.05 % Tween (TBST) for 30 min at room temperature and incubated with primary antibodies at 4 °C overnight. The following primary antibodies were used for western blotting: mouse anti-Flag (1:1,000, Sigma clone M2), rat anti-tubulin (1:5,000, Abcam clone YL1/2), mouse anti-Myc Tag (1:1,000, Cell Signaling 2276), anti-dicer (1:1,000) (Sinkkonen et al., 2010), rabbit anti-leo1 (1:1,000, Abgent AP1978A). Following incubation with corresponding HRP-conjugated secondary antibodies, signal was visualized using Immobilon Western Chemiluminiscent HRP Substrate (Merck Millipore). To detect biotinylated proteins, after transfer, membranes were blocked in 2 % bovine serum albumin (BSA) in TBST for 30 min and incubated with HRP-conjugated streptavidin (Strep-HRP) (Sigma S2438) diluted 1:10,000 in 2 % BSA-TBST for 30 min at room temperature.

**Co-immunoprecipitation experiments**

*Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 12E (AR50) or the corresponding parental cell line (cMB063) from one 10cm confluent dish were trypsinized, collected in media, washed in PBS and lysed for 30 min at 4°C in 500µl whole cell lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM KCl. 1.5 mM $MgCl_2$, 0.5% NP-40) supplemented with 1X protease inhibitor cocktail (Roche), 50 units benzonase and 10 µg RNase A. Lysates were clarified by centrifugation at 16,000g for 15 min and incubated with 30 µl M-280 Streptavidin Dynabeads (Invitrogen 11206D) for 3 hours at 4°C. The beads were washed 3 times in wash buffer (10 mM Tris-HCl pH 7.4, 150 mM KCl. 1.5 mM $MgCl_2$, 0.5% NP-40), resuspended in SDS loading buffer, boiled for 5min at 4°C and resolved by SDS-PAGE together with 5% of the input samples. Proteins were transferred semi-dry to a polyvinylidene fluoride (PVDF) membrane, blocked in 5 % non-fat milk in TBS+0.05 % Tween (TBST) for 30 min at room temperature and incubated with primary antibodies at 4 °C overnight. The following primary antibodies were used for western blotting: mouse anti-Flag (1:1,000, Sigma clone M2), rabbit anti-Paf1 (1:1,000, Abcam ab181120), rabbit anti-leo1 (1:1,000, Abgent AP1978A), rabbit anti-Ctr9 (1:1,1000, Abcam ab84487). Following incubation with corresponding HRP-conjugated secondary antibodies, signal was visualized using Immobilon Western Chemiluminiscent HRP Substrate (Merck Millipore).

**Affinity purification for LC–MS/MS**

For *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ single step streptavidin purification, two confluent 10cm dishes seeded with equal number cells of the tagged line or the corresponding parental cell line were harvested by trypsinization, washed twice in PBS and lysed for 1 hour at 4°C in 500µl whole cell lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM KCl. 2.5 mM $MgCl_2$, 0.5% NP-40) supplemented with 1X protease inhibitor cocktail (Roche), 50 units benzonase and 10 µg RNase A. Lysates were clarified by centrifugation at 16,000g for 10 min and incubated with 20 µl M-280 Streptavidin Dynabeads (Invitrogen 11206D) for 3 hours at 4°C. After that, the beads were washed four times in wash buffer (10 mM Tris-HCl pH 7.4, 150 mM KCl. 2.5 mM $MgCl_2$, 0.1% NP-40) and two times in wash buffer without NP-40. For mass spectrometry analysis, captured proteins were digested with trypsin directly on the streptavidin beads. *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ single step streptavidin purification was essentially carried out in the same way with the following exceptions: the cells were lysed in the absence of RNase and benzonase, in buffer containing 10mM Hepes pH 7.5, 150 mM KCl. 1.5 mM $MgCl_2$, 0.5% NP-40 supplemented with 1X PIC cocktail (Roche). Every affinity purification experiment contained three separate technical replicates for each cell line.

**Mass spectrometry analysis**

Peptides generated by trypsin digestion (see 'Affinity purification for LC–MS/MS') were analyzed by mass spectrometry as described in the "Methods" section of the manuscript related to Chapter 2 of this thesis. The full manuscript can be found in the Appendix.

**List of primers**

| Gene | Primer sequence | Orientation | Target | Application | mb Number |
|---|---|---|---|---|---|
| ***Leo1*** | GACGGTGAGGAGGAGCAAGAC | Forward | Exon2 | RT-qPCR | 8094 |
| | GTCTGCTTCGGAATCCGATGG | Reverse | Exon3 | RT-qPCR | 8095 |
| ***Leo1*** | CTGCCCAACTTCCTCAGTGTAG | Forward | Exon5 | RT-qPCR | 8098 |
| | CCTCTTCATCCAGCATCTCCTCATC | Reverse | Exon6 | RT-qPCR | 8099 |
| ***Tbp*** | TGCTGTTGGTGATTGTTGGT | Forward | Exon8 | RT-qPCR | 4325 |
| | AACTGGCTTGTGTGGGAAAG | Reverse | Exon8 | RT-qPCR | 4326 |
| ***Anks3*** | GACTGCACTCTTCCACTGTAC | Forward | Exon4-Exon5 | RT-qPCR | 9270 |
| | TGCTTCCATCAAAGGAGTGTATC | Reverse | Exon4-Exon5 | RT-qPCR | 9271 |
| ***Anks3*** | GAACCTGAGGCCTTGCAA | Forward | Exon1-Exon2 | RT-qPCR | 9272 |
| | GGACATCCAGCTCCTCTG | Reverse | Exon1-Exon2 | RT-qPCR | 9273 |
| ***Anks3*** | GCAGCCAAATGTTCATAGGGG | Forward | Intron4 | ChIP-qPCR | 9139 |
| | GTTCTGAGTGCATGAGAGCCT | Reverse | Intron4 | ChIP-qPCR | 9140 |
| ***Anks3*** | GGTTAGCAGAGAGGAGTTAGGC | Forward | Intron4 | ChIP-qPCR | 9137 |
| | TACCAACTGAACTGCCCAGC | Reverse | Intron4 | ChIP-qPCR | 9138 |
| ***Optn*** | TCCGAAATCAAGATGGAGCAG | Forward | Exon9-Exon10 | RT-qPCR | 9266 |
| | TCTGCCTGTTGCTTGGTTAG | Reverse | Exon9-Exon10 | RT-qPCR | 9267 |
| ***Optn*** | GAAGGTGTGGGAAGAGTGAG | Forward | 3′ UTR | RT-qPCR | 9268 |
| | GCAAGGTGTTTACTAAACTGGTG | Reverse | 3′ UTR | RT-qPCR | 9269 |
| ***Optn*** | TTGACACACTGTACCCCTCCT | Forward | 3′ UTR | ChIP-qPCR | 9125 |
| | CCATTCTGGCTGACTGCTCA | Reverse | 3′ UTR | ChIP-qPCR | 9126 |
| ***Radd23a*** | GGCAAGTGATTCAACAGAACC | Forward | Exon7-Exon8 | RT-qPCR | 9274 |
| | GGAGGCTCATTCAACATCTGG | Reverse | Exon7-Exon8 | RT-qPCR | 9275 |
| ***Radd23a/ Gadd45gip*** | GAGGAAGGAGGGTTGTCCTTG | Forward | 3′ UTR | qPCR | 9147 |
| | CTTAACAGCCCAGCCTACTCA | Reverse | 3′ UTR | qPCR | 9148 |
| ***Gadd45gip*** | CAAGAATGGTACCCGAGCTTAG | Forward | Exon1-Exon2 | RT-qPCR | 9276 |
| | AATCATTTGTGGCATCTTGGC | Reverse | Exon1-Exon2 | RT-qPCR | 9276 |
| ***MuERV*** | CAGAGTGATGAGAGACAGATGGA | Forward | Upstream of *MuERV* on chr1:114 | ChIP-qPCR | 8738 |
| ***MuERV*** | GTTGACAACCAGGAATAGCCAC | Reverse | 5′end of *MuERVL* LTR (MT2) | ChIP-qPCR | 8740 |
| ***Gapdh*** | CTCTGCTCCTCCCTGTTCC | Forward | Promoter | ChIP-qPCR | 9674 |

| | | | | | |
|---|---|---|---|---|---|
| ***Gapdh*** | TCCCTAGACCCGTACAGTGC | Reverse | Promoter | ChIP-qPCR | 9674 |
| ***Anks3*** | ACCAACACTAGTCATACCCACCTTCACACCCA | Forward | Promoter, contains SpeI site | Cloning | 9444 |
| | ACCAACGACGTCCTTGCAAGGCCTCAGGTTCC | Reverse | Exon1, contains AatII site | Cloning | 9445 |
| ***Anks3*** | ACCAACGACGTCCTTCTGGAGAGTGGAGCCA | Forward | Exon4, contains AatII site | Cloning | 9446 |
| | ACCAACGGCGCGCCGGAGTGTATCCATATACAGGCTCC | Reverse | Exon5, contains AscI site | Cloning | 9447 |
| ***Anks3 reporter*** | GTCACCTTCCCGGGATGCTC | Forward | Promoter-Exon1 | ChIP-qPCR | 12357 |
| | CAATCACCTCACGTTGGCATT | Reverse | Exon4 | ChIP-qPCR | 9264 |
| ***Anks3 reporter*** | GAACCTGAGGCCTTGCAA | Forward | Exon1 | ChIP-qPCR | 9272 |
| | GCCCACCCCTATGAACATTTG | Reverse | Intron4 | ChIP-qPCR | 12353 |
| ***Wdr61*** | CTGTCTGGCCTTTAGCGCTTCATCCTTGGTTAAGGAAATGgactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataaggctGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAggtcctACCAACCAGGTAAGGTCTGCAGTGAGAGGGAAAGCTTCAG | ssDNA ultramer | encodes a 3xFLAG-AviTag, flanked by 5′ and 3′ *Wdr61* homology arms | N-terminal endogenous gene tagging | 9656 |
| ***Skiv2l*** | AGTTTCCGGGCTGCTCGGGAGCTGCCGCGGCTCCAGGATGgactataaggaccacgacggagactacaaggatcatgatattgattacaaagacgatgacgataaggctGGCCTGAACGACATCTTCGAGGCTCAGAAAATCGAATGGCACGAAggtcctATGGAGACGGAGCGACTCGGTGAGGGGGAGGGGAGGAGAG | ssDNA ultramer | encodes a 3xFLAG-AviTag, flanked by 5′ and 3′ *Skiv2l* homology arms | N-terminal endogenous gene tagging | 10000 |

**List of TALENs**

| TALEN | Sequence | TALEN units | Source |
|---|---|---|---|
| HB-SBE B-globin locus | CCAGATTTGGTCACAGTT | HD1 HD2 NI3 NN4 NI5 NG6 NG7 NG8 NN9 NN10 NG1 HD2 NI3 HD4 NI5 NN6 NG7 LR-NG | Matyas Flemr, Bühler Lab |
| HB-SBK B-globin locus | GGAGGACAACTTTCCT | NN1 NN2 NI3 NN4 NN5 NI6 HD7 NI8 NI9 HD10 NG1 NG2 NG3 HD4 HD5 LR-NG | Matyas Flemr, Bühler Lab |
| SBE WDR61 | AGCGCTTCATCCTTGGTT | NI1 NN2 N*3 NN4 HD5 NG6 NG7 HD8 NI9 NG10 HD1 HD2 NG3 NG4 NN5 NN6 NG7 LR-NG | This work |
| SBK WDR61 | CTCACTGCAGACCTTACCT | HD1 NG2 HD3 NI4 HD5 NG6 NN7 HD8 NI9 NN10 NI1 HD2 HD3 NG4 NG5 NI6 HD7 HD8 LR-NG | This work |
| SBE SKIV2L | GCTCGGGAGCTGCCGCGG | NN1 HD2 NG3 N*4 NN5 NN6 NN7 NI8 NN9 HD10 NG1 NN2 HD3 N*4 NN5 N*6 NN7 LR-NN | This work |
| SBK SKIV2L | CTCACCGAGTCGCTCCGT | HD1 NG2 HD3 NI4 HD5 N*6 NN7 NI8 NN9 NG10 N*1 NN2 HD3 NG4 HD5 N*6 NN7 LR-NG | This work |

**List of cell lines**

| Identifier | Genotype | Clone | Source | Stable Integration |
|---|---|---|---|---|
| cMB053 | *Rosa26* $^{BirA\text{-}V52/-}$ | | Ostapcuk et al., 2018 | |
| cMB063 | *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ | | Ostapcuk et al., 2018 | |
| cMB068 | *Rosa26*$^{Cre\text{-}ERT2\text{-}5/-}$ *Leo1*$^{fl/fl}$ | Clone 5C | Alex Tuck, Bühler Lab | |
| cMB536 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ | Clone A1 | This work | DicerO |
| cMB537 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ | Clone 6F | This work | DicerO |
| cMB079 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ | Clone 3D | Alice Wenger, Bühler Lab | |
| cMB081 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ | Clone 6D | Alice Wenger, Bühler Lab | |
| cMB538 | *Rosa26* $^{BirA\text{-}V52/\ BirA\text{-}V5}$ *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | Clone 1G | This work | |
| cMB539 | *Rosa26* $^{BirA\text{-}V52/\ BirA\text{-}V5}$ *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | Clone 2G | This work | |
| cMB540 | *Rosa26* $^{BirA\text{-}V52/\ BirA\text{-}V5}$ *Wdr61*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | Clone 10H | This work | |
| cMB331 | *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | Clone 8F | This work | |
| cMB541 | *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | Clone 12E | This work | |
| cMB542 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ WT Dicer | Clone A2 | This work | WT Dicer |
| cMB543 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 2E | This work | DicerO; Anks3-Neon-Green reporter |
| cMB544 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 3B | This work | DicerO; Anks3-Neon-Green reporter |
| cMB545 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 7E | This work | DicerO; Anks3-Neon-Green reporter |
| cMB546 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 8F | This work | DicerO; Anks3-Neon-Green reporter |
| cMB547 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 9H | This work | DicerO; Anks3-Neon-Green reporter |
| cMB548 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 11G | This work | DicerO; Anks3-Neon-Green reporter |

| | | | | |
|---|---|---|---|---|
| cMB549 | *Rosa26* $^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ Dicer$^{O}$ Anks3-NeonGreen | Clone 12A | This work | DicerO; Anks3-Neon-Green reporter |
| cMB550 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3-NeonGreen | Clone 3A | This work | Anks3-Neon-Green reporter |
| cMB551 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3-NeonGreen | Clone 11F | This work | Anks3-Neon-Green reporter |
| cMB552 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3-NeonGreen | Clone 4D | This work | Anks3-Neon-Green reporter |
| cMB553 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3-NeonGreen | Clone 4F | This work | Anks3-Neon-Green reporter |
| cMB554 | *Rosa26*$^{Cre\text{-}ERT2/-}$ *Leo1*$^{fl/fl}$ *Dicer*$^{-/-}$ Anks3-NeonGreen | Clone 8A | This work | Anks3-Neon-Green reporter |

# Chapter 2

## INTRODUCTION

### 1. Cytoplasmic mRNA degradation

Eukaryotic mRNAs begin their life cycle in the nucleus where they are transcribed, capped, polyadenylated and spliced and then exported to the cytoplasm to serve as templates for translation into protein before they are finally degraded. Although regulation of gene expression is often mainly associated with the control of transcriptional events, RNA decay plays a key role in shaping the cellular proteome during normal conditions and in response to changing environmental cues (Bresson et al., 2017; Garneau et al., 2007; Pérez-Ortín et al., 2013). RNA degradation also supports the fidelity of gene expression through specialized cytoplasmic mRNA quality control pathways, which eliminate aberrant mRNAs that can generate potentially toxic protein products (Shoemaker and Green, 2012). The 5′ cap structure and 3′ poly(A) tail of eukaryotic mRNAs are two key features that confer stability by shielding the transcript ends from the cytoplasmic exonucleases. With the exception of non-polyadenylated histone mRNAs, transcript decay in the cytoplasm is typically initiated by gradual shortening of the poly-A tail and removal of the 5′ m7G cap (Figure 8) (Łabno et al., 2016a; Parker, 2012). At this point transcripts can become subject to degradation in the 5′-3′ direction by Xrn1, a conserved, highly processive exoribonuclease with a strong affinity towards substrates with 5′ monophosphates (Łabno et al., 2016a; Parker, 2012). Alternatively, mRNAs can be degraded in the 3′-5′ direction by the cytoplasmic exosome and its cofactor, the RNA helicase Ski complex (Łabno et al., 2016a; Parker, 2012). Dysregulation of RNA decay is implicated in several pathological conditions, underscoring the importance of this process in maintaining cellular homeostasis. For instance, reduced expression of Xrn1 is associated with osteosarcoma (Pashler et al., 2016), and mutations in exosome subunits have been linked to cancer and severe neurological disorders (Morton et al., 2018; Robinson et al., 2015). Furthermore, loss of function of the Ski complex causes tricho-hepato-enteric syndrome, a congenital bowl disorder that also affects multiple other organs (Fabre et al., 2012; Hartley et al., 2010).

RNA decay has been extensively studied in *S. cerevisiae* and several pioneering studies in this organism have been crucial in elucidating key mechanistic details of this process (Parker, 2012). The current model for RNA decay suggests that Xrn1 is the major cytoplasmic exonuclease, while the Ski-exosome complex is thought to operate redundantly with the 5′-3′ degradation pathway and make a more significant contribution in eliminating aberrant mRNAs (Łabno et al., 2016a; Parker, 2012). However, the precise

endogenous substrates of the cytoplasmic Ski-exosome pathway and its global function in RNA decay remain poorly defined particularly in mammalian cells. Thus, it is still an open question whether on a global scale, the two cytoplasmic pathways preferentially target certain transcripts, for instance based on specific RNA features. Furthermore, the presence of additional decay factors and pathways in higher eukaryotes (Łabno et al., 2016a) suggests an added level of complexity and highlights the importance of examining the process in other organisms.

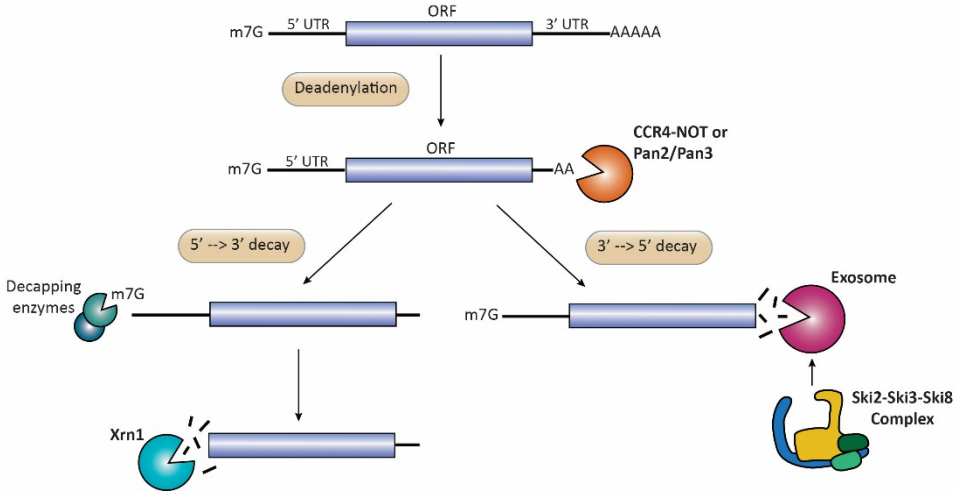



**Figure 8. General mRNA degradation pathways in the cytoplasm**
Degradation of cytoplasmic mRNA is typically initiated by shortening of the 3′ poly(A) tail catalyzed by the major deadenylase complexes Ccr4-NOT and Pan2/Pan3, followed by removal of the 5′ cap structure. At this point transcripts can be degraded in the 5′-3′ direction by the exoribonuclease Xrn1, or in the 3′-5′ direction by exosome and its co-factor the helicase Ski complex. Adapted with permission from (Garneau et al., 2007).

It is becoming increasingly clear that RNA decay is extensively coupled to other stages of the transcript life-cycle. In the cytoplasm in particular, a significant crosstalk exists between the translation and RNA degradation machineries. This is particularly evident for the cytoplasmic RNA surveillance pathways, which selectively degrade mRNAs containing aberrantly terminating ribosomes or stalled translation elongation complexes (Shoemaker and Green, 2012). However, accumulating evidence over the past decade suggests that the overlap between translation and RNA degradation is more widespread and not simply limited to RNA surveillance. In support of this idea, Xrn1 has been suggested to degrade mRNAs co-translationally (Hu et al., 2009; Pelechano et al., 2015), and recent structures of the yeast Ski complex and Xrn1 bound to the ribosome have established a direct physical link between the translation and RNA degradation machineries (Schmidt et al., 2016; Tesina et al., 2019). These observations raise interesting questions. First, is the physical association between the two processes conserved in higher eukaryotes?

Second, to what extent does each decay pathway interact with translation and what is the functional relevance of this interaction? This introduction focuses on several key aspects of yeast and mammalian cytoplasmic mRNA decay pathways and their links to translation.

### 1.1. mRNA Deadenylation

The poly(A) tails of eukaryotic mRNAs range in size from about 90 nucleotides in yeast to 200-250 in mammals and are coated by poly(A) binding protein (Pab1/PABPC1) (Eckmann et al., 2011). The poly(A) RNP is one of the most dynamic features of mature transcripts and plays a key role in controlling the stability of an mRNA (Jalkanen et al., 2014). Deadenylation of the poly(A) tail is considered the first and rate-limiting step in mRNA decay and is important for stimulating decapping and subsequent degradation in both 5′-3′ and 3′-5′ directions (Chen and Shyu, 2011; Parker, 2012).

Two conserved complexes, Ccr4-Not and Pan2-Pan3, are the major deadenylases that catalyze shortening of the poly(A) RNP (Wahle and Winkler, 2013). The predominant route for cytoplasmic deadenylation is thought to be through the Ccr4-Not complex which contains two catalytic subunits, Ccr4 and Caf1/Pop2 that act as 3′-5′ poly(A)-specific exoribonucleases (Tucker et al., 2002; Tucker et al., 2001). The catalytic activity of the Pan2-Pan3 complex lies in Pan2, which belongs to the RNaseD exoribonuclease family (Boeck et al., 1996; Uchida et al., 2004). Early studies in yeast have demonstrated that the two deadenylase complexes are differentially regulated by Pab1. While Pab1 appears to stimulate the activity of Pan2-Pan3 (Boeck et al., 1996), it has an inhibitory effect on the Ccr4-Not complex (Tucker et al., 2002). Thus, poly(A) tails bound by Pab1 would be protected from the activity of Ccr4-Not, but susceptible to degradation by Pan2-Pan3. This has become the basis of the widely-accepted bi-phasic model of mRNA deadenylation in both yeast and mammalian cells, which suggests that the two complexes act in a temporal manner. In the first phase, Pan2-Pan3 catalyzes initial trimming of the poly(A) tail, while subsequent deadenylation is carried out by the Ccr4-Not complex (Tucker et al., 2001; Yamashita et al., 2005).

Recent mechanistic studies in yeast and human cells have provided further details that clarify certain aspects of deadenylation. Contrary to previous observations, *in vitro* studies using components from fission yeast, demonstrate that Pab1 actually accelerates deadenylation through Ccr4-Not, by stimulating Ccr4, while inhibiting Caf1 (Webster et al., 2018). Consequently, the two catalytic subunits are proposed to have complementary roles, with Ccr4 displacing bound Pab1 and allowing the exposed poly(A) tail to be degraded by Caf1 (Webster et al., 2018). These findings are consistent with results obtained from reconstituted systems using human components (Yi et al., 2018), indicating that the mechanism of deadenylation by Ccr4-Not is conserved from yeast to humans.

Re-examining the role of the Pan2-Pan3 complex in human cells also suggests that certain aspects of the bi-phasic model need to be reevaluated. Unlike Ccr4-Not, perturbation of Pan2-Pan3 had only a minor impact on transcript abundance in HeLa cells, suggesting that the initial trimming of the poly(A) tail by Pan2-Pan3 is not essential or rate-limiting for deadenylation and decay (Yi et al., 2018). Thus, while Ccr4-Not seems to be the major deadenylase acting during bulk cytoplasmic decay of polyadenylated mRNAs, the functional role of initial poly(A) shortening by Pan2-Pan3 remains unclear (Yi et al., 2018).

### 1.2. mRNA Decapping

Deadenylated mRNAs are subject to decapping, which renders the transcripts susceptible to degradation in the 5′-3′ direction by the exoribonuclease Xrn1. Removal of the 5′ cap structure is mainly carried out by a conserved Nudix family hydrolase, Dcp2, which cleaves between the α - and β -phosphate of the cap to release m7GDP and 5′-end monophosphate RNA (Grudzien-Nogalska and Kiledjian, 2017). Dcp2 forms a complex with Dcp1, which enhances the intrinsic catalytic activity of the hydrolase and is required for decapping *in vivo* (Beelman et al., 1996; Chang et al., 2014a). Structural data from *S. pombe* indicate that Dcp1 stimulates the activity of Dcp2, by locking it in a particular closed, more active conformation. Interestingly, in mammalian cells this interaction surface is not well conserved, explaining why an additional factor, the scaffold protein EDC4, is required to bridge the Dcp1-Dcp2 interaction in higher eukaryotes (Chang et al., 2014a; She et al., 2008).

Multiple other activators known as enhancers of decapping (EDCs) function to stimulate the decapping machinery. Pat1, Lsm1-7, Dhh1 and Edc3 are among the factors conserved from yeast to human that promote decapping during general mRNA turnover (Grudzien-Nogalska and Kiledjian, 2017). Pat1 associates with Lsm1-7 to form a complex which preferentially binds to deadenylated mRNAs and enhances the interaction of Dcp2 with the transcript (Chowdhury et al., 2007; Chowdhury and Tharun, 2008; Nissan et al., 2010; Tharun and Parker, 2001). Pat1 has also been shown to interact with the 5′-3′ exoribonuclease Xrn1, thus establishing the Lsm1-7/Pat1 complex as an important link between deadenylation, decapping and mRNA degradation (Nissan et al., 2010).

Apart from Dcp2, recent studies in mammalian cells suggest that additional enzymes are capable of removing the cap structure of mRNAs. Dcp2 is not ubiquitously expressed in all mouse tissues and appears to be developmentally regulated (Song et al., 2010). Furthermore, exogenous RNA substrates could still be decapped in Dcp2-deficinet MEF cells, suggesting the presence of other enzymes that could act redundantly (Song et al., 2010). In fact, besides Dcp2, at least one other yeast and two mammalian proteins have also been shown to catalyze decapping (Grudzien-Nogalska et al., 2016; Jiao et al., 2010; Song et al., 2013; Song et al., 2010). Since the human genome encodes 22 Nudix proteins, future research is likely to reveal

additional decapping enzymes functioning *in vivo* (McLennan, 2006). This is another example of redundancy associated with RNA decay and highlights the complexity of this process particularly in higher eukaryotes.

## 1.3. Uridylation and its role in mRNA degradation

Although poly(A) tails of mature mRNAs are probably the best studied modification with respect to transcript stability, non-templated addition of other nucleotides has also been reported to affect the fate of RNA molecules. Much attention has been given to the extension of mRNA 3′-ends with short stretches of uridine residues, as research over the past fifteen years has established an important role for this modification in promoting cytoplasmic decay (De Almeida et al., 2018; Łabno et al., 2016a). Uridylation was first described in mouse and plants, where miRNA-directed cleavage products were found to contain short stretches of non-templated uridine residues at their 3′ ends, which was correlated with decapping and shortening of the 5′ cleavage product (Shen and Goodman, 2004). Subsequently, U-tails were also shown to be important for the turnover of non-polyadenylated histone mRNAs (Mullen and Marzluff, 2008). Uridylation is catalyzed by terminal uridylyl transferases (TUTases) or poly(U) polymerases (PUPs) and it is now clear that this modification is pervasive and conserved among all eukaryotes with the notable exception of *S. cerevisiae* (De Almeida et al., 2018).

Initial evidence that uridylation plays an important role in regulating the stability of cytoplasmic RNAs came from studies of the non-canonical nucleotidyltransferase Cid1 in *S. pombe*. Cid1 exhibits a robust poly(U) polymerase activity *in vitro* and was shown to catalyze the addition of U-tails to several polyadenylated mRNAs *in vivo* (Rissland et al., 2007; Rissland and Norbury, 2009). Cid1-dependent uridylation further promotes decapping of these poly(A)$^+$ mRNAs through the Lsm1–7 complex (Rissland and Norbury, 2009). These observations are consistent with earlier findings that describe the ability of the Lsm1–7/Pat1 complex to preferentially bind not only 3′ oligo(A) tails, but also 3′ oligo(U) stretches (Chowdhury et al., 2007). Interestingly, loss of Cid1 caused an accumulation of decay intermediates lacking 5′ cap structures, but with seemingly intact poly(A) tails (Rissland and Norbury, 2009). This suggests that in *S. pombe*, deadenylation was not a prerequisite for decapping, as this step could be bypassed by tagging the poly(A) tail with a stretch of uridines (Rissland and Norbury, 2009). Thus, deadenylation and uridylation were shown to have overlapping functions, suggesting that in *S. pombe* the two pathways operate in parallel, rather than sequentially (Rissland and Norbury, 2009).

The human ortholog of Cid1, ZCCHC6/TUT7, was also shown to exhibit poly(U) polymerase activity *in vitro* (Rissland et al., 2007) and an independent study demonstrated that in mammalian cell extracts, short 3′-end oligo(U) tracts added to a particular RNA promoted decapping through the Lsm1–7/Pat1 complex

(Song and Kiledjian, 2007), which is in line with observations from fission yeast. Although these early findings from *S. pombe* and mammalian cells demonstrated the conservation of uridylation and its impact on RNA degradation, it was unclear whether this process targets mRNAs on a global scale. The development of deep-sequencing approaches such as TAIL-seq, which provides information about poly(A) tail length and simultaneously detects 3′ end modifications (Chang et al., 2014b), were crucial for identifying the transcriptome-wide effects of uridylation. Recent applications of this technique in human cells revealed that approximately 85% of mRNAs contained 3′ U-tails at a frequency of higher than 1% suggesting that uridylation is pervasive (Chang et al., 2014b). Another study showed that human TUT4 and TUT7 act redundantly to catalyze mRNAs' uridylation both *in vitro* and *in vivo*, and loss of these factors resulted in a global increase in mRNA half-lives (Lim et al., 2014). Conversely, depletion of major cytoplasmic decay factors such as Xrn1 or the exosome caused an accumulation of U-tailed transcripts, suggesting that uridylation has a general role in facilitating RNA degradation (Figure 9) (Lim et al., 2014). Interestingly, unlike *S. pombe*, where uridylation seems to be independent of deadenylation, in human cells, U-tails were specifically added to deadenylated transcripts (Lim et al., 2014).

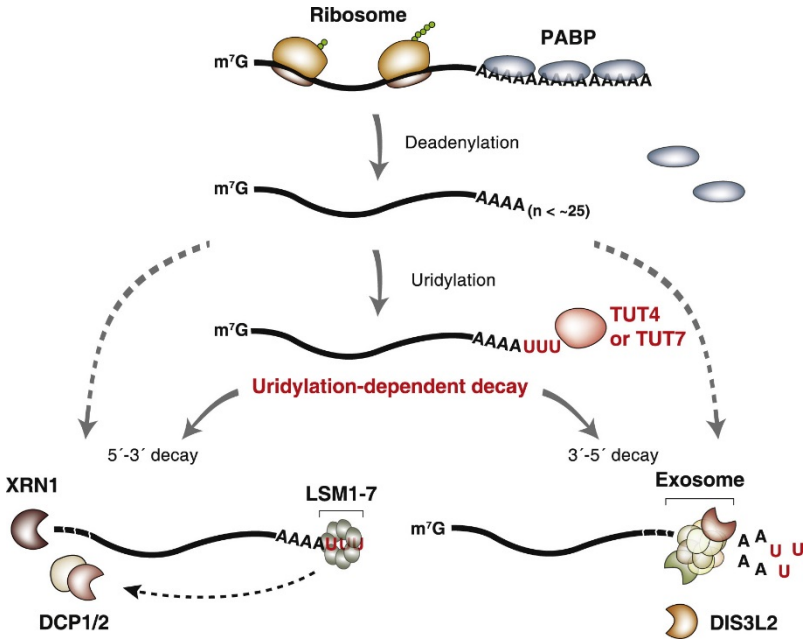



**Figure 9. Uridylation promotes mRNA degradation.**
TUT4 and TUT7 function redundantly to uridylate deadenylated mRNAs. U-tails can be recognized by the exosome or the exoribonuclease Dis3l2 to initiate 3′-5′ degradation of the transcript. The Lsm1-7 complex also binds U-tails and promotes decapping to allow degradation in the 5′-3′ direction by the exoribonuclease Xrn1. Adapted with permission from (Lim et al., 2014).

The same study also detected TUT4/7-dependent terminal oligo(U) stretches on histone mRNAs as well as other truncated decay intermediates lacking poly(A) tails (Lim et al., 2014). This suggests that besides its important role in bulk RNA turnover, uridylation may also facilitate the decay of RNA fragments generated by an internal endonucleolytic cleavage, which is often associated with RNA surveillance.

## 1.4. Cytoplasmic mRNA degradation by the 5′-3′ exoribonuclease Xrn1

Degradation of RNAs in the 5′-3′ direction by the processive exoribonuclease Xrn1 constitutes the major route for cytoplasmic mRNA decay (Parker, 2012). Besides targeting mRNAs that have been decapped after deadenylation, Xrn1 also degrades the byproducts of endonucleolytic cleavage such as those generated during cytoplasmic RNA surveillance (Jones et al., 2012). Initial evidence that Xrn1 is the predominant pathway for degradation of cytoplasmic mRNAs came from studies in *S. cerevisiae*, showing that blocking decay in the 5′-3′ direction caused perturbation in the normal rates of RNA turnover and was associated with more severe growth phenotypes compared to inactivation of the 3′-5′ pathway (Parker, 2012).

Consistent with its central role in cytoplasmic RNA decay, Xrn1 is highly conserved in all eukaryotes (Jones et al., 2012). The catalytic core of the enzyme is formed by two conserved N-terminal regions, CR1 and CR2, which coordinate a Mg2 + ion, necessary for catalysis (Chang et al., 2011; Jinek et al., 2011). Structural and biochemical studies have provided further explanation for the affinity of Xrn1 towards single-stranded substrates bearing a 5′ monophosphate (Jinek et al., 2011). The catalytic core of the exonuclease contains a conserved pocket of four basic residues which recognize a 5′ monophosphate, but exclude larger groups through steric hindrance (Jinek et al., 2011). This explains why Xrn1 cannot degrade the 5′ cap structure, thus perhaps ensuring that only transcripts that have already been committed to degradation after decapping or cleavage are targeted by the enzyme (Jinek et al., 2011). The high processivity of Xrn1 is achieved by recognition of the 5′ nucleotide of the RNA, which stimulates substrate translocation after each round of hydrolysis (Jinek et al., 2011). As the active channel is wide enough to accommodate only single stranded RNA, substrate translocation allows Xrn1 to combine hydrolysis with ATP-independent melting of more structured RNAs (Jinek et al., 2011).

Xrn1-mediated RNA decay has also been linked to other stages of the transcript life cycle. For instance, two studies implicate Xrn1 in a transcript buffering mechanism that couples RNA synthesis to degradation (Haimovich et al., 2013; Sun et al., 2013). More recently, Xrn1 has also been shown to associate with the ribosome during translation-coupled RNA decay (Tesina et al., 2019). These studies suggest that a significant crosstalk exists between RNA decay and other stages of gene expression.

## 1.5. Cytoplasmic 3′-5′ mRNA degradation by the exosome and the Ski complex

Cytoplasmic mRNAs can also be targeted for degradation in the 3′-5′ direction by the multi-subunit exosome complex and its cofactors (Parker, 2012). The exosome was discovered about 20 years ago in *S. cerevisiae* and since then orthologs have been found in a range of eukaryotes (Mitchell et al., 1997). This ribonuclease complex localizes to both the cytoplasm and nucleus where it targets various RNA species for degradation, processing, and surveillance, thus making a major contribution to RNA metabolism. The nuclear exosome is involved in the quality control of mRNAs, the maturation of stable non-coding RNAs, such as rRNA, snRNA and snoRNAs, and degrades the by-products of pervasive transcription, including cryptic unstable transcripts (CUTs) in yeast and promoter upstream transcripts (PROMPTS) in humans (Chlebowski et al., 2013; Kilchert et al., 2016). The cytoplasmic exosome functions during general mRNA turnover and has also been shown to degrade some highly unstable mRNAs containing AU-rich elements in their 3′ UTRs (Chen et al., 2001; Chlebowski et al., 2013). Transcripts targeted by the cytoplasmic RNA quality control pathways and RNAi are also exosome substrates (Chlebowski et al., 2013; Orban and Izaurralde, 2005).

The core exosome complex contains 9 catalytically inactive subunits, arranged in a two-layer, ring-like structure around a central channel which is wide enough to accommodate single-stranded RNA (Chlebowski et al., 2013). The first layer of the ring contains 6 proteins (Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and Mtr3) homologous to the RNase PH nucleases found in bacteria (Liu et al., 2006; Mitchell et al., 1997). The second layer, referred to as the ‘cap’, is formed by three RNA binding proteins, Rrp4, Rrp40 and Csl4 (Liu et al., 2006). The 9-subunit core serves as a binding platform for the ribonucleases Rrp6 and Dis3 (Rrp44), which provide enzymatic activity to the complex (Chlebowski et al., 2013). Rrp6 is a distributive 3′-5′ exonuclease of the DEDD family of nucleases and is restricted to the nucleus in both yeast and mammalian cells (Burkard and Butler, 2000; Januszyk et al., 2011). Dis3 is homologous to the bacterial RNase II/R and has both endonuclease and processive 3′-5′ exonuclease activities (Dziembowski et al., 2007; Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). In *S. cerevisiae*, Dis3 functions both in the cytoplasm and the nucleus. Human genomes encode two additional homologs of Dis3, Dis3L and Dis3L2 (Staals et al., 2010; Tomecki et al., 2010). Unlike *S. cerevisiae*, human Dis3 is mainly localized to the nucleus and Dis3L is restricted to the cytoplasmic exosome (Staals et al., 2010; Tomecki et al., 2010). Interestingly, while all human Dis3 homologs are processive 3′-5′ exonucleases, only the nuclear Dis3 has retained its endonuclease activity (Staals et al., 2010; Tomecki et al., 2010). Dis3L2 does not seem to associate with the exosome and instead functions separately as an alternative 3′-5′ degradation pathway in the cytoplasm (Lubas et al., 2013; Malecki et al., 2013). Biochemical and structural data indicate that RNA substrates can thread through the central channel of the catalytically inert core to access Dis3 (Bonneau et

al., 2009), which sits on the bottom of the channel. Alternatively, RNAs can be directly degraded by Rrp6 (Makino et al., 2015), which interacts with the top ('cap') of the core ring.

Multiple auxiliary factors and protein complexes associate with the exosome to regulate its access and recruitment to a wide range of substrates (Zinder and Lima, 2017). In mammalian cells, the nuclear exosome functions in the context of at least three different adapter complexes that facilitate RNA targeting (Fasken et al., 2011; Lubas et al., 2011; Meola et al., 2016). A central component of these complexes is the DEVH-box RNA helicase Mtr4/SKIV2L2, which unwinds substrates and funnels them into the narrow central channel of the exosome core (Weick et al., 2018). Similarly, the activity of the cytoplasmic exosome is coupled to its co-factor, the Ski complex, containing the DEVH-box RNA helicase Ski2/SKIV2L (Anderson and Parker, 1998; Halbach et al., 2013; Halbach et al., 2012). In *S. cerevisiase*, a separate protein Ski7, bridges the interaction between the cytoplasmic exosome and the Ski complex (Araki et al., 2001). The N-terminal portion of yeast Ski7 is required for binding to the Ski complex and exosome (Araki et al., 2001), while the C-terminal region has sequence similarity to translational GTPase domains (Guydosh and Green, 2014; Kowalinski et al., 2015), such as the one found in its closest paralog, Hbs1, functioning to release stalled ribosomes (Shoemaker et al., 2010). However, it was shown that whereas Ski7 is able to bind GTP, it is not a GTP-hydrolyzing enzyme (Kowalinski et al., 2015). Recent structural and biochemical studies in yeast also suggest that Ski7 is a constitutive component of the cytoplasmic exosome, which occupies similar surfaces on the core complex as Rrp6 (Kowalinski et al., 2016).

Until recently, it was unclear how the mammalian Ski co-factor interacts with the exosome, since Ski7 has no apparent orthologs outside fungi (Marshall et al., 2013). However, its closest paralog, Hbs1 is conserved across eukaryotes and two independent groups recently reported that a short splice variant of human HBS1L (HBS1L3) is the Ski7-like mammalian protein that can interact with both the exosome and the Ski complex, thus confirming the evolutionary conservation of this co-factor (Kalisiak et al., 2017; Kowalinski et al., 2016). Since yeast Ski7 and human HBS1L3 share only a sparse sequence similarity, it would be interesting to examine how each of these factors recruits the Ski complex to the exosome (Kowalinski et al., 2016).

The Ski complex is the only known cytoplasmic co-factor of the exosome. It is evolutionary conserved and assists all exosome-mediated RNA degradation pathways, including general RNA turnover and surveillance (Anderson and Parker, 1998; Araki et al., 2001; Mitchell and Tollervey, 2003; van Hoof et al., 2002; van Hoof et al., 2000). Recently, the yeast Ski complex has also been shown to physically interact with the 80S ribosome (Schmidt et al., 2016), suggesting that it is involved in translation-coupled decay, similar to yeast Xrn1.The Ski complex consists of the tetratricopeptide protein Ski3/TTC37, the RNA-dependent RNA helicase Ski2/SKIV2L, and two copies of the beta-propeller Ski8/WDR61 (Halbach et al., 2013; Synowsky and Heck, 2008). Ski2 is structurally and functionally related to its nuclear counterpart Mtr4, although

biochemically Ski2 seems to have a broader substrate specificity than Mtr4 and has affinity for both single-stranded and double-stranded RNAs (Halbach et al., 2012). The N-terminal portion of Ski2 interacts with Ski3 and Ski8, while the C-terminal region harbors the helicase activity (Halbach et al., 2013; Halbach et al., 2012). The catalytic core of Ski2 is formed by two RecA domains (RecA1 and RecA2), which fold into a globular structure that mediate substrate binding and ATP hydrolysis (Halbach et al., 2012). RecA2 has a conserved unwinding β-hairpin region, which is a typical feature of the DExH-box helicases (Halbach et al., 2012). The C-terminal region also contains an RNA-binding 'insertion' domain which forms an elongated structure protruding from the globular core (Halbach et al., 2012).

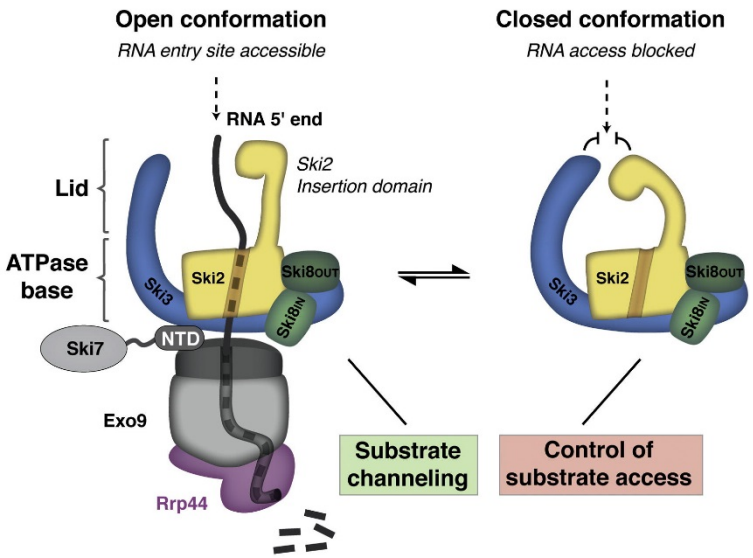



**Figure 10. Model for RNA channeling into the exosome-Ski complex.**
Substrate access into the ATPase core of the Ski complex is controlled by interactions between the N-terminal arm of Ski3 and the insertion domain of Ski2 which restrict the RNA binding and ATPase activities of the complex. Upon activation, RNA is threaded through the Ski2 helicase and into the core exosome (Exo9) channel where it reaches the catalytic exoribonuclease Rrp44 (Dis3). Ski7 is proposed to bridge the interaction between the two complexes. Adapted with permission from (Halbach et al., 2013).

Although the catalytic activity of the complex is provided entirely through the helicase Ski2, structural and biochemical studies of the yeast Ski complex have revealed important functions for Ski3 and Ski8 (Halbach et al., 2013). In addition to the N-terminal region of Ski2, Ski3 and both copies of Ski8 make contacts with the globular helicase core as well, which stabilizes RNA binding to Ski2 and extends its RNA binding

surface (Halbach et al., 2013). More importantly, the N-terminal arm of Ski3 and the insertion domain of Ski2 form a lid-like structure, which controls the ATPase activity of the complex, possibly by regulating substrate access to the catalytic core (Figure 10) (Halbach et al., 2013). Consistent with this, the ATPase rate of Ski2 in isolation increases about 5-fold compared to the activity of the whole complex, while the ATP binding properties remain unchanged (Halbach et al., 2013). Thus, besides contributing to the overall structural integrity of the complex, Ski3 and Ski8 also modulate its RNA binding and ATPase properties (Halbach et al., 2013). Furthermore, Ski3 and Ski8 are important for linking the complex to the exosome, since Ski7 does not interact directly with Ski2 (Wang et al., 2005). Biochemical experiments suggest that the Ski complex can directly thread single-stranded RNA substrates into the exosome channel, combining the helicase and nuclease activities of the two machineries (Figure 10) (Halbach et al., 2013).

Our understanding of Ski-exosome-mediated RNA degradation in the cytoplasm comes primarily from studies in *S. cerevisiae*. Details about the regulation and function of the Ski complex in higher eukaryotes are limited, although abnormalities of the exosome and its cytoplasmic co-factor have been implicated in human Mendelian diseases (Fabre and Badens, 2014). Current models for cytoplasmic decay suggest that the yeast Ski complex operates redundantly with Xrn1. However, loss of function of Xrn1 has been associated with specific phenotypes in *Drosophila* and *C. elegans* (Grima et al., 2008; Newbury and Woollard, 2004; Zabolotskaya et al., 2008) and similarly, perturbation of the Ski complex causes the rare congenital disorder trico-hepato-enteric syndrome (Fabre et al., 2012), suggesting that at least *in vivo*, the two decay pathways are not completely redundant in multicellular organisms.

### 1.6. Degradation of non-polyadenylated histone mRNAs

Replication-dependent histone mRNAs are the only known examples of metazoan mature coding transcripts which are not polyadenylated. Consequently, regulation of histone mRNA turnover differs significantly from the canonical degradation pathways of poly(A)+ transcripts. Instead of a poly(A) tail, histone mRNAs contain a conserved 26-bp stem-loop (SL) structure within their 3′ UTRs, which is bound by a stem loop binding protein (SLBP) (Battle and Doudna, 2001; Williams and Marzluff, 1995). This feature regulates all aspects of histone mRNA metabolism, including processing, export, translation and RNA turnover (Marzluff and Koreski, 2017). The expression of histones is regulated at multiple levels and is tightly coupled to the cell cycle to ensure proper packaging of DNA into chromatin. While histone mRNAs accumulate at the onset of DNA synthesis, they are rapidly degraded at the end of S-phase (Harris et al., 1991; Whitfield et al., 2000). This balance between the amount of histones and DNA synthesis is important, since artificially elevated levels of histones have been linked to chromosome loss and genome instability (Meeks-Wagner and Hartwell, 1986).

The conserved stem loop structure at the 3′end of histone transcripts acts as a *cis* regulatory sequence element that is both necessary and sufficient to induce histone mRNA decay at the end of S-phase (Pandey and Marzluff, 1987). The stem loop forms a tight ternary complex with the SLBP and the 3′-5′ exonuclease Eri1 (Tan et al., 2013), which has been shown to be important for degradation of the SL structure (Hoefig et al., 2013). Eri1 is conserved from fission yeast to humans and has an N-terminal SAP domain important for RNA binding, followed by a DEDD family nuclease domain (Dominski et al., 2003; Yang et al., 2006). The SLBP occupies the 5′ side of the SL, while Eri1 interacts with the loop, the 3′ arm and the 3′ flanking sequence of the SL (Tan et al., 2013). SLBP and Eri1 do not contact each other directly in the ternary complex and have been shown to recognize the overall structure of the stem loop, rather than specific sequences (Tan et al., 2013).

The initial trigger of histone mRNA degradation is thought to be uridylation of their 3′ ends (Mullen and Marzluff, 2008). Several TUTases have been implicated in this process, including TUT1, TUT3, TUT4 and more recently TUT7 (Lackey et al., 2016; Mullen and Marzluff, 2008; Schmidt et al., 2011). Addition of oligo(U) tails has been suggested to promote degradation of the stem loop by Eri1, consistent with the observed accumulation of uridylated full length histone mRNAs in Eri1 knock down cells (Hoefig et al., 2013). The removal of the stem loop by Eri1 would then allow the transcripts to be degraded in the 3′-5′ direction by the exosome (Slevin et al., 2014). Furthermore, U-tailed histone mRNAs are also bound by the Lsm1-7 complex, which has been suggested to repress translation and mediate decapping, allowing the transcript to be degraded in the 5′-3′ direction (Hoefig et al., 2013; Lyons et al., 2014; Mullen and Marzluff, 2008). Consistent with this, knock down of Lsm1, Dcp2, or Xrn1 reduces the rates of histone mRNA degradation in HeLa cells (Mullen and Marzluff, 2008). Interestingly, in vitro experiments demonstrated that Eri1 can efficiently degrade single-stranded RNA and overhangs, but is strongly inhibited by double-stranded RNA structures (Yang et al., 2006), which raises the question as to how Eri1 degrades through the stem loop *in vivo*. Several observations suggest that the RNA helicase Ufp1 might be involved in this process. In addition to its essential role in NMD, Upf1 is also important for histone mRNA degradation (Kaygun and Marzluff, 2005a; Mullen and Marzluff, 2008) and has been shown to interact with the SLBP, in an RNA-independent manner (Hoefig et al., 2013). Upf1 was also found to co-immunoprecipitate with Eri1 and proteins of the Lsm1-7 complex (Hoefig et al., 2013). It is therefore possible that melting of the stem-loop structure by Upf1 facilitates Eri1-mediated degradation. However, a precise molecular mechanism underlying the proposed role for Upf1 is still missing. Furthermore, since Eri1 has been shown to associate with the Lsm1-7 complex in an RNA-independent manner (Hoefig et al., 2013), it remains unclear if and how these factors cooperate during histone mRNA degradation. The initial events that trigger biochemical changes at the 3′ stem loop RNP, recruitment of TUTases and subsequent uridylation are also not very well understood. It has also been suggested that Eri1 may function to oppose the constant addition

of uridines by TUTases, thus facilitating efficient translation, since stretches of only five uridines have been shown to significantly interfere with histone translation (Hoefig and Heissmeyer, 2014). In this scenario, the continuous removal of histone U-tails by Eri1 would be counteracted at the end of S-phase, for instance by an unknown factor.

Although histone mRNAs can be degraded both in the 5′-3′ and 3′-5′ direction (Mullen and Marzluff, 2008), the relative contribution of each pathway is unclear. When introduced into S-phase HeLa cells, luciferase reporter constructs containing histone 3′ stem loops and Dcp2-resistant caps were degraded much slower than reporters with normal caps (Su et al., 2013). Similarly, knock down of Dcp2 delayed degradation of reporters with normal cleavable caps, suggesting that 5′-3′ degradation preceded by decapping is the predominant pathway for histone degradation (Su et al., 2013). Other studies, however, report that knock down of exosome subunits has a larger stabilizing effect on histone mRNAs, arguing for a more significant contribution of the 3′-5′ decay pathway (Mullen and Marzluff, 2008; Slevin et al., 2014). Consistent with this, many histone mRNA decay intermediates with partially degraded stem loops or ORFs were found to be capped (Slevin et al., 2014). Interestingly, the same study also reports that knock down of the exosome-independent 3′-5′ exonuclease Dis3L2 does not affect histone mRNA degradation, although Dis3L2 has been implicated in this pathway in another study (Łabno et al., 2016b).

## 2. Cytoplasmic mRNA quality control pathways

In addition to the general decay pathways initiated by deadenylation and decapping, cells have evolved specialized mechanisms that monitor the quality of cytoplasmic mRNAs and selectively eliminate aberrant transcripts. These mechanisms comprise three mRNA surveillance pathways: nonsense-mediated decay (NMD), no-stop decay (NSD) and no-go decay (NGD), which target transcripts with premature stop codons, no in-frame stop codons, and mRNAs containing stalled ribosomes, respectively (Shoemaker and Green, 2012). A unifying feature of these pathways is their close link to the ribosome, whereby detection of an aberrant translation event is required to trigger mRNA degradation (Shoemaker and Green, 2012).

### 2.1. Nonsense-mediated decay (NMD)

Translation of mRNAs containing premature stop codons (PTCs) could produce truncated protein products with potentially toxic or dominant-negative properties. Therefore, timely detection and elimination of such transcripts through the NMD pathway is important for cellular homeostasis. PTC-containing mRNAs could arise from random nonsense or frameshift DNA mutations, or pre-mRNA processing events, which could produce spliced isoforms containing additional stop codons within the main ORF (Lykke-Andersen and

Jensen, 2015). NMD can also result from the translation of alternative PTC-containing ORFs within the full-length mature transcript. For instance, it has been estimated that approximately 35-50% of upstream ORFs (uORFs) can trigger degradation of the host mRNA through NMD (Somers et al., 2013).

The initial step in the NMD pathway requires detection of a prematurely terminating ribosome, which leads to the assembly of a multi-protein complex that promotes accelerated mRNA degradation (Lykke-Andersen and Jensen, 2015). Central to the formation of this complex is the helicase Upf1, which is a conserved primary NMD factor important for both substrate recognition and degradation (Kim and Maquat, 2019; Lykke-Andersen and Jensen, 2015). Key to the function of Upf1 in NMD are its RNA-dependent ATPase and 3′-5′ helicase activities (Franks et al., 2010; Kurosaki et al., 2014; Weng et al., 1996) as well as phosphorylation of its N- and C-terminus (Ohnishi et al., 2003; Yamashita et al., 2001). Upf1 has several distinct domains including a disordered N-terminal region, followed by a cysteine- and histidine-rich (CH) sequence and helicase domain, and a C-terminal region containing serines and glutamines (SQ) (Kim and Maquat, 2019). The CH region of Upf1 represses its ATPase activity by folding onto the helicase domain (Chamieh et al., 2008; Fiorini et al., 2013). Interaction between Upf1 and two other core NMD factors, Upf2 and Upf3, induces a large conformational change in the CH domain, which stimulates the ATPase and helicase activity of Upf1 (Chakrabarti et al., 2011; Chamieh et al., 2008). In addition, the kinase Smg1 phosphorylates the [S/T]Q motifs in the N- and C-terminal regions of Upf1 which triggers recruitment of downstream NMD factors such as the endonuclease Smg6 or the Smg5–Smg7 complex (Chakrabarti et al., 2014; Okada-Katsuhata et al., 2012). Smg6 catalyzes endonucleolytic cleavage close to the aberrantly-terminating ribosome, producing RNA fragments that can be degraded by the major cytoplasmic exonucleases (Eberle et al., 2009; Huntzinger et al., 2008; Lykke-Andersen et al., 2014; Schmidt et al., 2015). Furthermore, the Smg5–Smg7 dimer promotes substrate deadenylation and subsequent decapping through recruitment of the Ccr4-Not deadenylase complex (Loh et al., 2013; Unterholzner and Izaurralde, 2004). NMD can also promote cap removal more directly, via recruitment of the general decapping complex or the decapping enhancer Pnrc2 (Cho et al., 2013; Cho et al., 2009; Lai et al., 2012).

The precise mechanism of target recognition that allows the NMD machinery to distinguish premature stop codons from normal translation termination events is not fully understood. Upf1 has been shown to associate with the 3′ UTRs of actively translated mRNAs and was also observed within the coding sequence after translation inhibition (Hogg and Goff, 2010; Hurt et al., 2013; Kurosaki and Maquat, 2013; Zund et al., 2013). How Upf1 can specifically target aberrantly terminating ribosomes is unclear, but it is generally thought that its direct interaction with the canonical translation termination factors eRF1 and eRF3 is the initial event triggering an NMD (Lykke-Andersen and Jensen, 2015). Stop codons within the A-site of the ribosome are recognized by the eRF1-eRF3 complex and subsequent GTP hydrolysis by eRF3 promotes

the release of the nascent polypeptide chain and dissociation of the ribosomal subunits (Dever and Green, 2012). Several studies have suggested that stop codons in close proximity to the poly(A) binding protein are less prone to NMD, possibly due to the ability of PABP to promote efficient termination and prevent the interaction between eRF3 and Upf1 (Eberle et al., 2008; Silva et al., 2008; Singh et al., 2008). According to this model short 3′UTRs would correlate with more efficient translation termination, which would counteract NMD. Conversely, longer 3′UTRs seem to predispose some mRNAs to NMD, possibly due to the increased distance between the stop codon and the poly(A) tail (Figure 11) (Buhler et al., 2006a; Eberle et al., 2008; Singh et al., 2008). An additional consequence of a longer 3′UTR might be higher non-specific Upf1 occupancy, which could increase the probability of forming an interaction between eRF3 and Upf1.

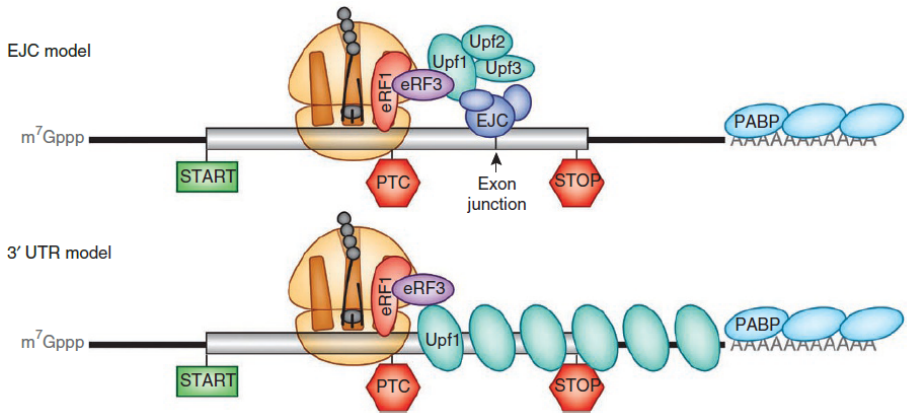



**Figure 11. Models for NMD substrate recognition**
In the EJC model, the presence of a premature stop codon (PTC) upstream of an exon-junction complex facilitates NMD. Upf1 interacts with eRF3 engaged with the terminating ribosome as well as Upf2 and Upf3 which are associated with the EJC. Upf2 and Upf3 cooperatively stimulate the ATPase and helicase activities of Upf1.
In the 3′ UTR model, due to the extended distance between the PTC and PABP, translation termination is inefficient which facilitates interaction between eRF3 and Upf1. In this model, longer 3′ UTRs would predispose transcripts to NMD, also possibly due to increased non-specific binding of Upf1. Adapted with permission from (Shoemaker and Green, 2012).

Multiple other studies support a role for the exon-junction complex (EJC) in NMD. EJCs are deposited 20-24 nucleotides upstream of each exon-exon junction during pre-mRNA splicing in the nucleus (Le Hir et al., 2000). Since authentic stop codons are expected to be located in the last exon, the presence of an EJC downstream of a stop codon represents an aberrant event favoring NMD (Figure 11) (Lykke-Andersen and Jensen, 2015). However, not all NMD events are dependent on EJCs, as illustrated by the ability of some non-spliced transcripts to undergo NMD (LeBlanc and Beemon, 2004; Rajavel and Neufeld, 2001) and the

fact that PTCs can trigger degradation even after the pioneer round of translation, when EJCs would no longer be present along the ORF (Rufener and Muhlemann, 2013).

## 2.2. Non-stop (NSD) and no-go decay (NGD)

Non-stop decay (NSD) and no-go decay (NGD) are cytoplasmic quality control mechanisms that target messages lacking an in-frame stop codon or mRNAs containing stalled translation elongation complexes, respectively (Doma and Parker, 2006; Frischmeyer et al., 2002; van Hoof et al., 2002). Non-stop substrates could potentially arise from truncated or cleaved mRNAs, where the translating ribosome would stall upon reaching the end of the template (Shoemaker and Green, 2012). Another class of non-stop targets includes transcripts lacking a termination codon, but containing a poly(A) tail (Shoemaker and Green, 2012). Such mRNAs could be formed for example by alternative polyadenylation which has been shown to be widespread and can occur within the ORF (Frischmeyer et al., 2002; Ozsolak et al., 2010). Translation of the poly(A) sequence in this case would also result in translational arrest (Arthur et al., 2015; Garzia et al., 2017; Guydosh and Green, 2017; Ikeuchi and Inada, 2016; Juszkiewicz and Hegde, 2017; Sundaramoorthy et al., 2017). The no-go pathway elicits degradation of transcripts containing ribosomes stalled at stable RNA structures, stretches of poly-basic codons or chemically damaged residues among others (Doma and Parker, 2006; Guydosh and Green, 2017; Ikeuchi et al., 2019; Simms et al., 2014; Winz et al., 2019). Although often referred to as two separate pathways, both NSD and NGD involve arrested translation elongation complexes that are resolved through the same mechanism (Figure 12). Another characteristic feature of the combined NSD/NGD pathway is an endonycleolytic cleavage upstream of the stalling sequence (D'Orazio et al., 2019; Doma and Parker, 2006; Guydosh and Green, 2017; Ikeuchi et al., 2019; Simms et al., 2017; Tsuboi et al., 2012). The resulting 5′ and 3′ RNA fragments can then be degraded by Xrn1 and the exosome. Subsequent rounds of surveillance-coupled endonycleolytic cleavage can also occur if an upstream translating ribosome stalls upon reaching the end of the cleaved 5′ RNA fragment. In support of this model, several studies report the presence of regularly-spaced cleavage events upstream of the stalling sequence (Guydosh and Green, 2017; Simms et al., 2017).

Two conserved factors, Dom34 (Pelota in mammals) and Hbs1, recognize and resolve stalled ribosome complexes during NSD/NGD (Guydosh and Green, 2014; Pisareva et al., 2011; Shoemaker et al., 2010; Shoemaker and Green, 2011; Tsuboi et al., 2012). Dom34 and Hbs1 are structural and functional homologs of the canonical translation termination factors, eRF1 and eRF3, and similarly form a heterodimer that can interact with the A-site of the ribosome (Becker et al., 2011; Chen et al., 2010a; Graille et al., 2008; Lee et al., 2007). Hbs1 is a GTPase with substantial sequence similarity to eRF3, and Dom34 is structurally related to eRF1 with two important distinguishing features. Unlike eRF1, Dom34 does not contain the conserved

GGQ motif that is necessary to catalyze peptide hydrolysis, and its N-terminal domain lacks the NIKS sequence motif involved in stop codon recognition (Chen et al., 2010a; Graille et al., 2008; Lee et al., 2007). As a result, the presence of a specific A-site codon is not a prerequisite for Dom34:Hbs1 binding, allowing the complex to function more broadly as a general rescue factor for a variety of translational stalls.

The Dom34:Hbs1:GTP complex can supposedly sense a vacant A-site on the ribosome during a prolonged translational arrest (Figure 12). Upon binding to Hbs1, Dom34 adopts a conformation more closely resembling a tRNA, which supposedly facilitates interaction with the A-site (Chen et al., 2010a). After GTP hydrolysis and Hbs1 release, the concerted action of Dom34/PELOTA and the recycling factor Rli1/ABCE1 dissociates the ribosomal subunits (Pisareva et al., 2011; Shoemaker and Green, 2011). Unlike eRF1 which also promotes peptide-tRNA hydrolysis, Dom34 releases a peptide-tRNA conjugate, which remains associated with the 60S subunit (Pisareva et al., 2011; Shoemaker et al., 2010; Shoemaker and Green, 2011). This unusual nascent polypeptide-tRNA-60S complex is recognized and processed by the ribosome quality control (RQC) pathway, which leads to nascent chain proteasomal degradation after ubiquitination by the E3 ubiquitin ligase Listerin (Joazeiro, 2019). Thus, mRNA surveillance is coupled to nascent chain quality control to eliminate all defective components of an aberrant translation event.

Although Dom34:Hbs1 does not exhibit preference for a specific A-site codon, the recycling activity of the yeast and mammalian complex has been shown to depend on the length of the mRNA downstream of the P-site (Pisareva et al., 2011; Shoemaker and Green, 2011). While the *S. cerevisiae* Dom34:Hbs1 complex seems more permissive to various 3′ mRNA lengths (Shoemaker and Green, 2011), the mammalian Pelota:Hbs1 has been shown to efficiently recycle stalled ribosomes if they contained only up to 9 mRNA nucleotides downstream of the P-site, while the recycling activity was completely inhibited by the presence of more than 12 nucleotides (Pisareva et al., 2011). This suggests that the two rescue factors would primarily function on stalled ribosome complexes containing very short or no mRNA overhangs. Consistent with this, Dom34/Pelota:Hbs1 has been shown to split empty 80S ribosomes (Pisareva et al., 2011; Shoemaker and Green, 2011; van den Elzen et al., 2014). Furthermore, studies in yeast have demonstrated that Dom34 primarily targets ribosomes that stall either on the edge of truncated mRNAs or close to the end of the 3′ UTR (Guydosh and Green, 2014). In contrast, Dom34 did not seem to act on various classes of stall-inducing sequences, such as poly-prolines that were located within the ORF (Guydosh and Green, 2014). Structural data suggest that this preference for arrested ribosomes with short 3′ mRNA overhangs might be due to Hbs1 whose N-terminus contacts the mRNA entry site and may survey the 3′ mRNA length (Becker et al., 2011). Consistent with this, ribosome splitting by Dom34 and the recycling factor Rli1 in yeast is not dependent on the 3′ mRNA length in the absence of Hbs1 (Shoemaker and Green, 2011). This preference for short 3′ mRNA overhangs raises the question of how/if Dom34:Hbs1 is able to act on the primary

stalling event, since an endonucleolytic cleavage that can generate a short 3′ mRNA end has been shown to occur upstream of the stall-inducing sequence. Regardless of the mechanism, Dom34/Pelota and Hbs1 can become important for ribosome rescue after secondary cleavage events upstream of the initial stall site (Guydosh and Green, 2017).

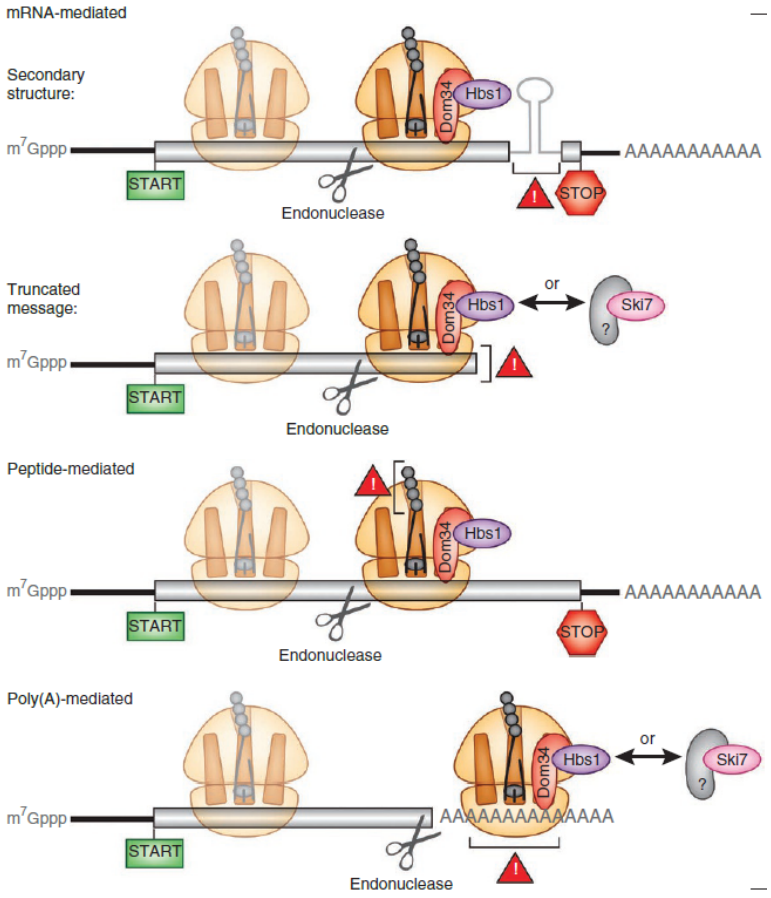



**Figure 12. Triggers of NGD and NSD**
Although NGD and NSD are commonly referred to as separate processes, both events converge on stalled translation elongation complexes and involve an endonucleolytic cleavage at the vicinity of the stalled ribosome. An empty ribosome A-site during translational arrest is recognized by the rescue factors Dom34 (Pelota in mammals) and Hbs1, which facilitate dissociation of the ribosome subunits. Potential triggers of ribosome stalling include secondary mRNA structures, translation of truncated messages or sequences that encode inhibitory peptides. Ribosomes translating mRNAs lacking an in-frame stop codon also arrest at the poly(A) tail and become substrates for Dom34:Hbs1. Adapted with permission from (Shoemaker and Green, 2012).

Recognition of arrested elongation complexes by Dom34:Hbs1 does not seem to be a prerequisite for mRNA cleavage, since ribosome stalls seem to trigger endonucleolysis even in the absence of Dom34:Hbs1

(Chen et al., 2010a; Guydosh and Green, 2017; Kuroha et al., 2010). As the identity of the endonuclease has remained obscure for a long time, a recent study in yeast provides strong evidence that Cue2 is the conserved enzyme that cleaves mRNAs during NGD/NSD (D'Orazio et al., 2019). Distinguishing features of Cue2 include two N-terminal CUE ubiquitin-binding domains and a C-terminal SMR (small MutS-related) hydrolase domain (D'Orazio et al., 2019). Although DNA nicking has been described as the canonical function of MutS-related SMR enzymes (Fukui and Kuramitsu, 2011), some proteins with SMR domains have been shown to function as endoribonucleases (Bhandari et al., 2011; Zhou et al., 2017), consistent with the SMR structural similarity to the bacterial RNase E (Fukui and Kuramitsu, 2011). NGD reporter cleavage was completely abolished by removal of the C-terminal SMR domain of Cue2 or by mutations in its catalytic residues (D'Orazio et al., 2019). Further analysis revealed that Cue2 cleaves mRNAs at the A-site of the trailing ribosome within a collided ribosome pair, in line with structural homology modelling, which places the catalytic domain of Cue2 at the A-site of the 40S subunit (D'Orazio et al., 2019). It was subsequently shown that ribosomes accumulating at cleaved sites were Dom34 substrates, further supporting a role for Cue2 in the NGD pathway (D'Orazio et al., 2019).

Putative Cue2 homologs exist in multiple organisms (D'Orazio et al., 2019), and another study in *C. elegans* identified the SMR-domain containing protein Nonu-1 as the endonuclease involved in NSD (Glover et al., 2019). Interestingly, the mammalian Cue2 homolog, N4BP2 contains a PNK domain (D'Orazio et al., 2019). Since an endonucleolytic cleavage event can generate a 3′ RNA fragment carrying a 5′ hydroxyl group, the kinase activity of N4BP2 could potentially phosphorylate the 3′ cleavage product to facilitate subsequent Xrn1-mediated degradation (D'Orazio et al., 2019).

## 3. Ribosome collisions as a platform for mRNA and protein quality control

As stalled translation elongation complexes are the underlying trigger of mRNA surveillance, an important question is how the cells initially recognize terminally arrested ribosomes to elicit degradation of the mRNA and the arrested nascent polypeptide chain. Genetic studies in *S. cerevisiae* identified the 40S ribosomal protein Asc1 (Rack1 in mammals) and the ubiquitin ligase Hel2 (Znf598 in mammals) as two important factors involved in the early events of ribosome-associated quality control pathways (Brandman et al., 2012; Kuroha et al., 2010). Deletion of Asc1 or Hel2 in yeast results in increased read-through of stall-inducing poly-basic sequences and production of full length reporter proteins, suggesting that these factors might be necessary for initial recognition of ribosome stalls (Brandman et al., 2012; Kuroha et al., 2010). The WD-repeat β-propeller protein Asc1/Rack1 is an integral component of the 40S ribosomal subunit located near the mRNA exit channel (Gerbasi et al., 2004; Nielsen et al., 2017). Besides its ribosome-associated functions, Asc1/Rack1 is also involved in different cell signaling pathways (Nielsen et al., 2017).

Hel2/Znf598 is a RING-type E3 ubiquitin ligase that has been shown to directly associate with translating ribosomes (Garzia et al., 2017; Singh et al., 2012; Winz et al., 2019).

Subsequent studies have demonstrated that regulatory ubiquitination of 40S ribosomal subunits mediated by Hel2/Znf598 and Asc1/Rack1 is the initial event that marks arrested ribosomes for surveillance (Garzia et al., 2017; Juszkiewicz and Hegde, 2017; Matsuo et al., 2017; Sitron et al., 2017; Sundaramoorthy et al., 2017; Winz et al., 2019). Translational stalling on a poly-arginine reporter in yeast has been shown to require Hel2-mediated ubiquitination of the 40S subunit uS10 (Rps20) at Lysine 6 and 8 (Matsuo et al., 2017). Subsequent ribosome splitting and degradation of the arrested nascent chain were also dependent on this modification (Matsuo et al., 2017). Conversely, deletion of Hel2 or disruption of its interaction with the 40S subunit, allowed translational read-through of the stall-inducing sequence, effectively abolishing induction of the ribosome quality control pathway (Matsuo et al., 2017; Sitron et al., 2017; Winz et al., 2019).

The mammalian ortholog of Hel2, Znf598, has also been shown to mediate site-specific ubiquitination of 40S subunits. More specifically, several studies demonstrated that ribosome stalling on poly-lysine reporters encoded by consecutive AAA-codons, was dependent on eS10 (Rps10) and uS10 (Rps20) ubiquitination by Znf598 (Garzia et al., 2017; Juszkiewicz and Hegde, 2017; Sundaramoorthy et al., 2017). Similar to the *Hel2*Δ phenotype, loss of Znf598 or mutations that render its 40S targets resistant to ubiquitination, promote increased read-through of the poly-lysine reporters and suppress induction of the ribosome quality control pathway (Garzia et al., 2017; Juszkiewicz and Hegde, 2017; Sundaramoorthy et al., 2017). Although Rack1 does not possess an enzymatic activity, it has been implicated in regulating ubiquitination of uS3 (Rps3), uS5 (Rps2) and uS10 (Rps20) (Sundaramoorthy et al., 2017). However, lack of Rack1 resulted in a less pronounced read-through phenotype compared to knockdown of Znf598, and the combined loss of both proteins did not have an additive effect, suggesting that Rack1 and Znf598 might have partially overlapping functions within the same pathway (Sundaramoorthy et al., 2017). A recent study suggests that yeast Asc1 helps recruit and retain Hel2 on the ribosome, providing a possible explanation for the observed Asc1/Rack1 role in ribosome stalling and ubiquitination (Winz et al., 2019).

Although 40S ubiquitination by Hel2/Znf598 and Asc1/Rack1 has been primarily studied in the context of ribosome stalling, subunit dissociation and degradation of the nascent polypeptide chain, these two proteins also facilitate endonucleolytic cleavage and degradation of NSD and NGD substrates (Ikeuchi and Inada, 2016; Ikeuchi et al., 2019; Winz et al., 2019). This suggests that Hel2/Znf598 and Asc1/Rack1 are early-acting factors that function prior to ribosome splitting and RNA decay. Given that the recently discovered NGD endonuclease Cue2 in yeast has ubiquitin binding domains, it is possible that Hel2-ubiquitinated sites might help recruit Cue2 to arrested ribosomes to initiate RNA decay (D'Orazio et al., 2019).

Since Hel2/Znf598 and Asc1/Rack1 have been shown to specifically target stalled translation complexes, it is unclear how they distinguish between temporary or programmed translational pauses versus pathologic terminally arrested ribosomes. Some forms of permanently stalled ribosomes such as those arriving at the end of a truncated template would contain an empty A-site, making them an optimal substrate for the ribosome rescue factors Dom34/Pelota-Hbs1 (Guydosh and Green, 2014). However, subjecting other forms of more temporary pauses to surveillance might be inappropriate, as they could represent normal events that facilitate proper folding of the nascent chain, recruitment of chaperones or localization of the mRNP (Collart and Weiss, 2019). A recent study suggests that individually paused 80S ribosomes can evade surveillance since they are not optimal targets for 40S ubiquitination (Juszkiewicz et al., 2018). Instead, Znf598 has been shown to specifically recognize and ubiquitinate a di-ribosome or disome structure, which arises when an elongating trailing ribosome collides with a slower or stalled leading ribosome (Juszkiewicz et al., 2018). This disome structure is defined by extensive 40S-40S interactions between the stalled and collided ribosomes which positions the mRNA channels of the two ribosomes in close proximity, forming a nuclease-resistant unit (Figure 13) (Juszkiewicz et al., 2018). Structural data indicate that the inter-ribosomal 40S-40S interactions form two distinct interfaces (Juszkiewicz et al., 2018). Interface 1 includes ribosomal subunits eS1 (Rps3A), uS11 (Rps14), eS26 (Rps26), and eS28 (Rps28), which surround the mRNA exit channel of the stalled ribosome (Figure 13) (Juszkiewicz et al., 2018). Interestingly, interface 2 juxtaposes Rack1 of the leading ribosome and uS3 (Rps3), uS10 (Rps20), and eS10 (Rps10) of the trailing ribosome, which were previously shown to be targeted for ubiquitination (Figure 13) (Juszkiewicz et al., 2018). This is consistent with a model where Znf598 specifically recognizes interface 2, which would explain why it preferentially associates with and ubiquitinates eS10 (Rps10) within disomes as well as higher order collided ribosome structures *in vitro* (Juszkiewicz et al., 2018). Similarly, inducing global ribosome collisions *in vivo* by subjecting HEK293 cells to sub-inhibitory concentrations of various translation elongation inhibitors coincided with Znf598 localization to polysomes and elevated levels of ubiquitinated eS10 (Rps10) (Juszkiewicz et al., 2018).

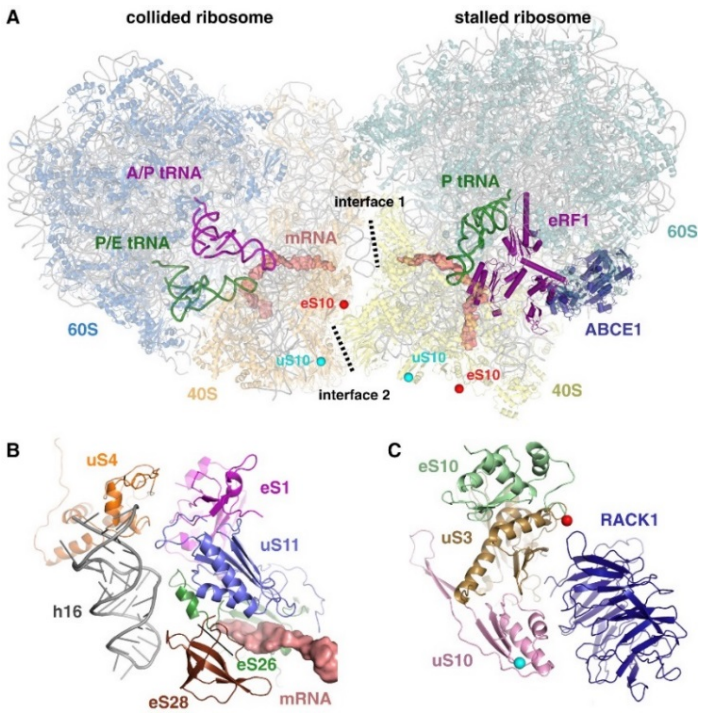



**Figure 13. Structure of the collided di-ribosome**
**A** The collided di-ribosome or disome structure is primarily formed by interactions between the 40S subunits of the stalled and collided ribosomes. The ubiquitination sites of uS10 (Rps20) and eS10 (Rps10) are indicated with cyan and red dots, respectively. **B** and **C** Closer view of the two interfaces formed by the 40S-40S interactions. Interface 1 shown in B is formed by eS1 (Rps1), uS11 (Rps14), eS26 (Rps26) and eS28 (Rps28) of the stalled ribosome and uS4 (Rps9) and h16 of the 18S rRNA of the collided ribosome. Interface 2 shown in C is formed by Rack1 of the stalled ribosome and uS3 (Rps3), uS10 (Rps20) and eS10 (Rps10) of the collided ribosome. Adapted with permission from (Juszkiewicz et al., 2018).

The collided di-ribosome conformation has also been shown to facilitate Hel2-mediated quality control pathways in yeast (Ikeuchi et al., 2019). Consistent with studies in mammalian cells, cryo-EM analysis of disomes isolated from *S. cerevisiae*, confirmed that this structure is formed primarily by interactions between the 40S subunits of the stalled and collided ribosomes, which bring Asc1 and all Hel2 substrates in close proximity (Ikeuchi et al., 2019). Furthermore, similar to Znf598, Hel2 was also shown to preferentially ubiquitinate disomes over monosomes (Ikeuchi et al., 2019). This was important not only for degradation of the nascent polypeptide chain, but also for mRNA cleavage within the disome unit during NGD, suggesting that the two events are coupled by Hel2-mediated disome ubiquitination (Ikeuchi et al., 2019). Another recent study using NGD reporters in yeast also suggests that ribosome collisions are essential for endonucleolytic cleavage and mRNA degradation during NGD (Simms et al., 2017). In this study, ribosome stalling stem-loop structures or poly-basic repeat sequences initiated robust NGD only

when they were located further than 105nt from the start codon (Simms et al., 2017). Polysome analysis subsequently showed that NGD-incompetent reporters harboring stalling sequences near the AUG codon were translated by only a few ribosomes, which reduces the likelihood of collisions and suggests that higher ribosome occupancy on the translated mRNA promotes endonucleolytic cleavage and robust NGD (Simms et al., 2017). Furthermore, consistent with previous observations, inducing global ribosome collisions *in vivo* triggered Hel2-mediated ubiquitination of the 40S subunit (Simms et al., 2017). Collectively, findings from studies in yeast and mammalian cells suggest that aberrant translation events are often defined by collided ribosomes which serve as platforms for recruitment of downstream quality control factors.

## 4. Co-translational mRNA degradation

RNA decay is coupled to other stages and events of the transcript lifecycle and a growing body of evidence points to a significant crosstalk between translation and mRNA degradation. For instance, efficient translation initiation and elongation have been generally associated with increased transcript stability. More recently, the differential decoding efficiency of the 61 codons, referred to as 'codon optimality' has also been shown to affect mRNA half-lives (Presnyak et al., 2015; Radhakrishnan et al., 2016). A transcriptome-wide study of mRNA decay rates in *S. cerevisiae* found a striking correlation between non-optimal codons and shorter mRNA half-lives (Presnyak et al., 2015). The same study also showed that altering codon optimality was sufficient to change transcript stability (Presnyak et al., 2015). Since decoding non-optimal codons slows down ribosome translocation, elongation rates dictated by codon usage emerge as a key determinant of transcript stability (Presnyak et al., 2015).

The extensive crosstalk between RNA decay and translation is best illustrated by the cytoplasmic mRNA surveillance pathways, which dispose of faulty transcripts co-translationally. Since substrate recognition depends on detection of aberrant translation termination or elongation events, inhibition of translation results in stabilization of surveillance targets (Shoemaker and Green, 2012). However, accumulating evidence suggests that co-translational degradation might be more widespread and not simply limited to the mRNA quality control pathways. Consistent with this model early studies in yeast and human cells found that the general decay factors Xrn1 and Dcp2 co-sediment with polysomal fractions (Mangus and Jacobson, 1999; Wang et al., 2002). Components of the decapping and deadenylation machinery were also shown to co-purify with ribosome complexes in *Drosophila* cells (Antic et al., 2015). Furthermore, decapped and deadenylated mRNA decay intermediates were associated with polysomes in yeast and *Drosophila* cells, suggesting that dissociation of ribosomes from the mRNA is not a prerequisite for RNA degradation (Antic et al., 2015; Hu et al., 2009).

Further evidence supporting pervasive co-translational decay during bulk mRNA degradation comes from systematic profiling of yeast polyadenylated mRNAs bearing a 5′ monophosphate (5P-seq), which mainly represent decapped degradation intermediates generated by the enzymatic activity of Xrn1 (Pelechano et al., 2015). Interestingly, the distribution of these 5′P molecules within the coding mRNA sequence was not random but instead displayed a pronounced, frame-specific 3-nucleotide periodicity reminiscent of the codon-wise translocation of the ribosome (Pelechano et al., 2015). This pattern was dependent on Xrn1 and was shown to be weaker around the beginning of the ORF, which is consistent with a model where after decapping, Xrn1 chases, catches up and closely follows the last translating ribosome (Pelechano et al., 2015).

Another recent study in human cells surveyed global mRNA degradation by sequencing the 3′ and 5′ ends of capped and polyadenylated decay intermediates, respectively (Ibrahim et al., 2018). Similar to previous findings in yeast, the authors observed a striking 3-nucleotide periodicity for both ends of decay intermediates mapping to coding sequences (Ibrahim et al., 2018). However, this ribosome-phased pattern was not affected by the knockdown of Xrn1 or certain exosome subunits (Ibrahim et al., 2018). Instead, by further examining the distribution of degradation fragments relative to ribosome protected footprints, the authors concluded that the 3-nucleotide periodicity must have been produced by repeated ribosome-associated endonucleolytic cleavage events (Ibrahim et al., 2018). Although the general cytoplasmic exonucleases did not seem directly involved in generating the 3-nucleotide pattern, Xrn1 and the exosome were suggested to trim and clear the ribosome-associated fragments downstream of the endonuclease (Ibrahim et al., 2018). Approximately two-thirds of the 50% most highly expressed mRNAs were estimated to be degraded through such co-translational process, suggesting that this is not a quality control pathway for aberrant transcripts, but a more general mechanism that targets canonical mRNAs engaged by ribosomes (Ibrahim et al., 2018).

Consistent with the co-translational mRNA decay model, *S. cerevisiae* Xrn1 was recently reported to directly and specifically associate with translating ribosomes (Tesina et al., 2019). The cryo-EM structure of the Xrn1-ribosome complex shows that the exonuclease interacts predominantly with the small 40S subunit, making direct contacts with the mRNA and ribosomal proteins surrounding the mRNA exit channel (Tesina et al., 2019). These interactions suggest that mRNA exiting the ribosome decoding center is funneled directly into the catalytic core of Xrn1 (Tesina et al., 2019). This tight coupling between translation and decay is proposed to further contribute to the processivity of the exonuclease, by facilitating substrate channeling and stimulating Xrn1 conformational changes throughout the catalytic cycle (Tesina et al., 2019). However, *S. cerevisiae* strains carrying an endogenous mutation that completely abolished the association of Xrn1 with polysomes were indistinguishable from WT strains, unlike *xrn1*Δ mutants which

exhibited a significant growth phenotype (Tesina et al., 2019). Perturbing the interaction between the exonuclease and the ribosome also had little effect on the degradation of a known Xrn1 substrate, although it is not clear how this mutation affects mRNA decay on a global scale (Tesina et al., 2019).

The 3′-5′ RNA degradation pathway was also shown to be physically linked to translation in a recent cryo-EM structure capturing the *S. cerevisiae* Ski complex associated with a translating 80S ribosome (Schmidt et al., 2016). The Ski complex was bound near the 40S mRNA entry tunnel presenting a short 3′ mRNA overhang that was threaded directly into the Ski2 helicase core (Schmidt et al., 2016). The lid-like structure formed by Ski3 and the insertion domain of Ski2, which autoinhibits the ATPase properties of the helicase, adopts a more open conformation upon binding to the 40S subunit, suggesting that direct coupling to the ribosome might stimulate the enzymatic activity of the Ski complex (Schmidt et al., 2016). The interaction between the two complexes was increased when the 80S ribosome contained 3′ mRNA overhangs between 10 and 20 nucleotides and was weakened with shorter or longer mRNA stretches (Schmidt et al., 2016), suggesting that the Ski complex might preferentially target ribosomes presenting certain RNA lengths downstream of the A-site. One potential source of such substrates could be co-translational endonucleolytic cleavage events (Schmidt et al., 2016). Targeted ribosomal profiling revealed that the Ski-ribosome complex was enriched over non-optimal codons suggested to promote decay through translational stalling and it is conceivable that an endonuclease acting on such transcripts could create an ‘access point’ for the Ski complex (Schmidt et al., 2016). Interestingly, contrary to previous models, the Ski complex was found to associate with 80S ribosomes stalled on a non-stop reporter even in the absence of Ski7 and Dom34, suggesting that the helicase complex functions upstream of these factors and likely recruits a preassembled Ski7-exosome complex to its substrates to initiate degradation (Schmidt et al., 2016). It is currently unclear whether the direct interaction between the ribosome and the main cytoplasmic decay factors is conserved in higher eukaryotes and what fraction of transcripts would be targeted for degradation co-translationally.

# RESULTS

## 1. Mammalian RNA decay pathways are highly specialized and widely linked to translation

Alex Charles Tuck[1,4], Aneliya Rankova[1,4], Alaaddin Bulak Arpat[3], Luz Angelica Liechti[3], Daniel Hess[1], Vytautas Iesmantavicius[1], Violeta Castelo-Szekely[3], David Gatfield[3] and Marc Bühler[1,2,5]

*Molecular Cell, 2020 (The full manuscript can be found in the Appendix)*

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
[2]University of Basel, Petersplatz 10, 4003 Basel, Switzerland
[3]Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland
[4]These authors contributed equally
[5]Lead contact

### 1.1. Summary

The prevailing model for cytoplasmic mRNA degradation, largely based on studies in *S. cerevisiae,* suggests that the 5′-3′ exonuclease Xrn1 is the predominant route for mRNA degradation, whereas the 3′-5′ Ski-exosome pathway is thought to function redundantly with Xrn1, and contribute more significantly to RNA surveillance (Parker, 2012). However, the precise endogenous targets of the two pathways are poorly defined in higher eukaryotes, including mammals. **Thus, it is still an open question whether on a global scale, certain mRNAs are preferentially degraded via one pathway, and if so, what factors might underlie such specificity.** Furthermore, cytoplasmic RNA decay is influenced by translation and studies over the past decade suggest that mRNAs are widely targeted for degradation while still engaged with actively translating ribosomes (Hu et al., 2009; Pelechano et al., 2015). Consistent with this, recent cryo-EM structures of the yeast Ski complex and Xrn1 bound to the ribosome have now established a direct physical link between the translation and RNA degradation machineries (Schmidt et al., 2016; Tesina et al., 2019)**. However, it is currently unknown whether the physical association between the two processes is conserved in higher eukaryotes. It also remains unclear to what extent each decay pathway interacts with translation, and what factors might influence this. Here, we set out to address these questions by examining the two major cytoplasmic decay pathways and their links to translation in mouse embryonic stem cells.**

We first performed crosslinking and analysis of cDNAs (CRAC) to identify the transcriptome-wide targets of Xrn1 and the Ski complex helicase Skiv2l in unperturbed cells (Figure 14A and 14B). Although some

transcripts were similarly occupied by Xrn1 and Skiv2l, for others binding of one factor seemed to dominate, suggesting that under normal conditions certain transcripts might be primarily degraded via one route (Figure 14C). Further focusing on 3′-5′ decay, we found that substrates highly bound by Skiv2l, were also upregulated upon Skiv2l knock out, indicating that the 3′-5′ degradation pathway is important for regulating the steady-state levels of this subset of mRNAs. Our CRAC data further revealed that both Xrn1 and Skiv2l crosslink to the small subunit of the ribosome, in a manner closely resembling the recent yeast cryo-EM structures, confirming that the physical coupling between translation and RNA degradation is conserved in mammals (Figure 14D). Interestingly, we observed major differences in the binding patterns of Xrn1 and Skiv2l along mRNAs. While Xrn1 was enriched in the 3′ UTRs, Skiv2l was exclusively bound to ribosome-occupied regions and was largely absent from 3′ UTRs (Figure 14E). These findings are surprising given that the activity of the yeast Ski complex is required for efficient degradation of 3′UTRs (Anderson and Parker, 1998) and previous studies have shown that Ski2 is bound to 3′ UTRs in yeast (Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013). Our data suggest that the mammalian 3′-5′ Ski-exosome pathway does not normally degrade full-length transcripts, and its activity is instead tightly coupled to co-translational RNA decay. Consistent with this, we show that Skiv2l is enriched at sites of abnormal translation, known to trigger mRNA surveillance, such as NMD substrates, upstream ORFs (uORFs) and other sequences encoding amino acid stretches that can cause ribosome stalling.

Given the preference of Skiv2l for certain substrates as revealed by our CRAC data, in combination with the striking correlation between Skiv2l mRNA binding and ribosome occupancy, we next wondered whether the Ski complex associates with specific adapter proteins that regulate its recruitment to ribosome-bound sequences. Using IP-MS experiments, we identified two factors, Focadhesin and the RNA binding protein Aven, as interaction partners of the mammalian Ski complex (Figure 15A and 15B). Interestingly, Aven and Skiv2l shared similar ribosome and mRNA binding sites, suggesting that they might function in the same pathway (Figure 15C and 15D). Loss of Aven led to marked accumulation of Skiv2l on most of its substrates (Figure 15E) and a substantial increase in Skiv2l CRAC reads containing non-templated 3′-end U-tails, indicating an increase in 3′-5′ decay.

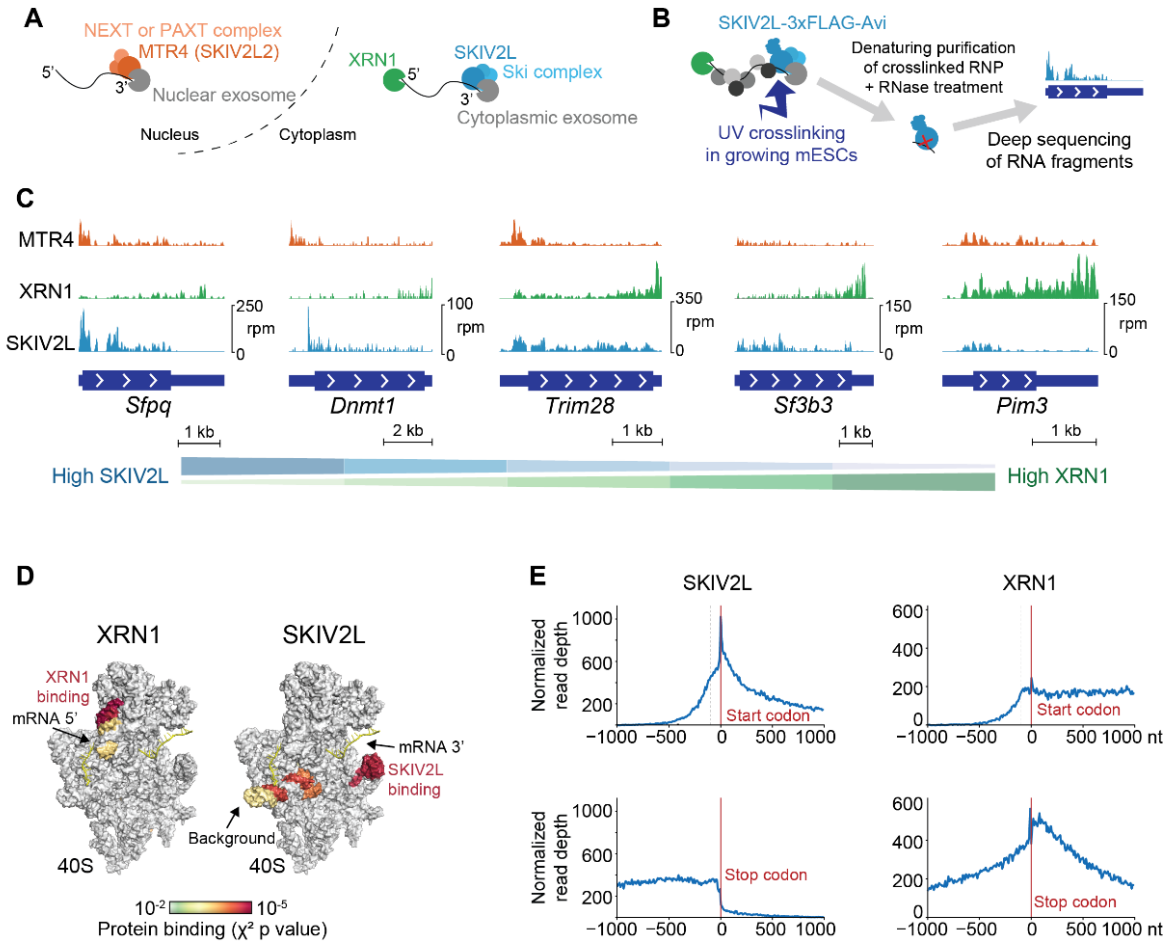



**Figure 14. mRNA binding profiles of Xrn1 and Skiv2l in mESC.**
**A** Major decay pathways in the nucleus and cytoplasm. **B** Schematic of the CRAC method. **C** CRAC coverage of Mtr4, Xrn1 and Skiv2l along individual mRNAs. **D** Xrn1 and Skiv2l CRAC signal on the ribosomal 40S subunit, based on the mouse 18S rRNA sequence and human structure. Significantly bound regions are colored by $\chi 2$ p-value. **E** Cumulative distribution of Skiv2l and Xrn1 CRAC reads around the start and stop codons of mRNAs. Adapted from (Tuck et al., 2020).

Since the distribution of Skiv2l across transcripts is tightly coupled to ribosomes, we hypothesized that the changes in Skiv2l binding in the absence of Aven could reflect altered ribosome occupancy. Indeed, global profiling of disomes, or collided di-ribosome pairs that arise at translation stall sites, revealed that in the absence of Aven, disome occupancy was significantly increased in transcripts gaining Skiv2l binding (Figure 16). These data suggest that Aven could act to facilitate translation and prevent sustained ribosome stalls that would otherwise trigger mRNA surveillance and Skiv2l recruitment. This is in line with a previously reported role for Aven in promoting efficient translation of highly structured mRNAs (Thandapani et al., 2015), and is further supported by our own data showing that Aven binding is enriched at sequences with high structural propensity.

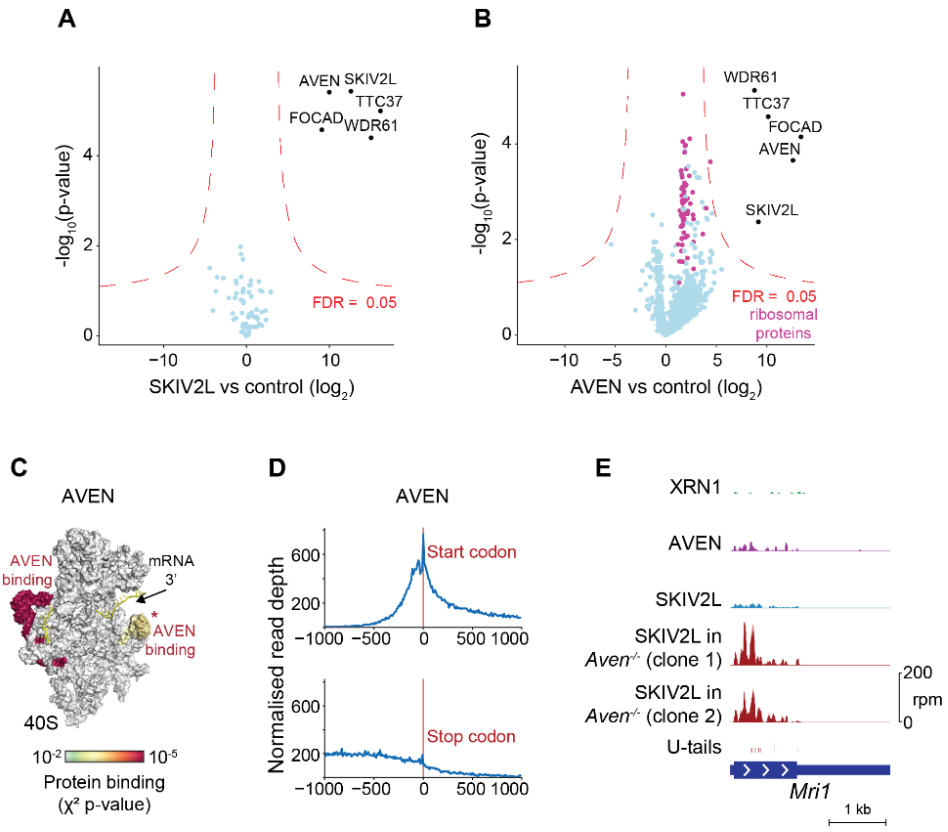



**Figure 15. The mammalian Ski complex interacts with Focadhesin (Focad) and Aven.**
**A** Mass spectrometry analysis (MS) of tandem FLAG-streptavidin purification of endogenously tagged 3xFLAG-Avi-Skiv2l. **B** MS of single strep streptavidin purification of endogenously tagged 3xFLAG-Avi-AVE. **C** Aven CRAC signal on the ribosome 40S subunit and **D** around the start and stop codons of mRNAs. **E** Example of an mRNA that gains Skiv2l binding (measured by CRAC) in the absence of Aven. Adapted from (Tuck et al., 2020).

Collectively, our findings are consistent with a model where Aven and Skv2l act in concert to limit aberrant translation, whereby Aven counteracts prolonged ribosome stalls and Skiv2l helps eliminate the mRNA if the stalls accumulate and cannot be resolved. Intriguingly, the Aven-Skiv2l pathway acts on a wide range of substrates, including mRNAs, uORFs and most surprisingly on small-ORF-containing RNAs derived from transcription of non-coding regions. Since Aven is conserved from *Drosophila* to humans, our work reveals a higher-eukaryote-specific mechanism that coordinates cytoplasmic 3′-5′ RNA decay.

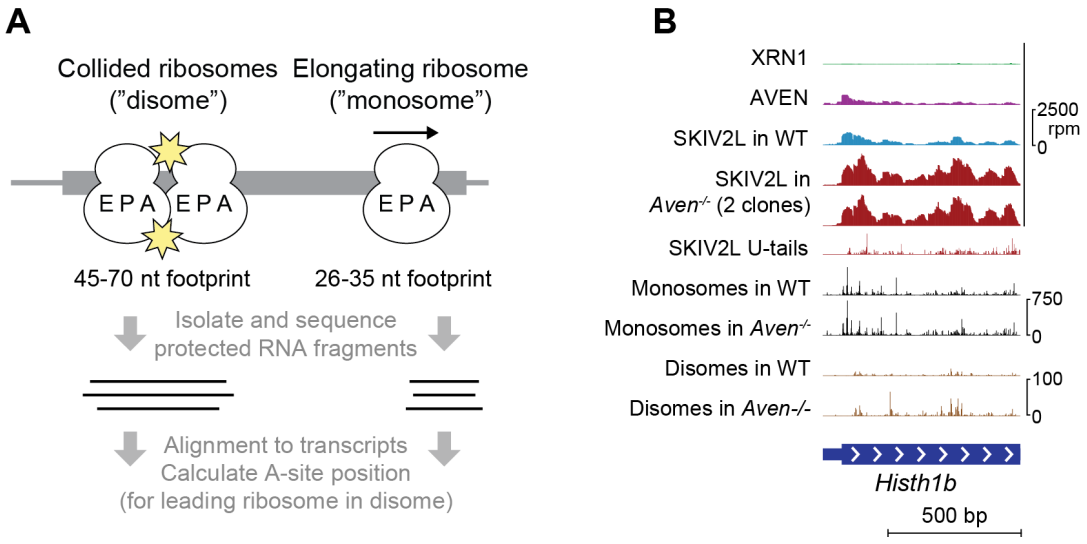



**Figure 16. Skiv2l binding and disome occupancy along mRNAs increase in the absence of Aven.**
**A** Schematic highlighting the differences between conventional monosome foot-printing and disome profiling which maps sites of collided ribosome pairs. **B** Example of an mRNA that gains substantial Skiv2l binding (measured by CRAC) and disome density upon Aven knock out. Adapted from (Tuck et al., 2020).

## 1.2. Contributions

This was a shared project between me and Alex Tuck, a postdoctoral researcher in the Bühler Lab. Both Alex and I created cell lines and performed CRAC and RNA-seq experiments. Genome-wide data analyses were done by Alex Tuck. I performed IP-MS experiments discovering the interaction between the mammalian Ski complex and Aven/Focadhesin. The IP-MS data were analyzed by Daniel Hess and Vytautas Iesmantavicius from the Protein Analysis Facility at the FMI. In addition, we had an excellent collaboration with the Gatfield Lab from the University of Lausanne who performed all ribosome profiling experiments. More specifically, monosome and disome profiling was done by Luz Liechti, and related analyses were performed by Alaaddin Arpat and Violeta Castelo-Szekely from the Gatfield Lab.

# DISCUSSION

## 1. Cytoplasmic RNA degradation is widely linked to translation

It is well established that cytoplasmic RNA decay can be influenced by translation. Although mRNA surveillance, which depends on ongoing translation, is one of the best examples demonstrating the close links between the ribosome and the decay machinery, accumulating evidence suggests that co-translational decay may be more widespread (Antic et al., 2015; Hu et al., 2009; Ibrahim et al., 2018; Pelechano et al., 2015). Recent cryo-EM structures of yeast Xrn1 and the Ski complex bound to the translating ribosome have further established a direct link between the two processes (Schmidt et al., 2016; Tesina et al., 2019). This raises the question of whether the physical interaction between translation and mRNA decay is conserved in higher eukaryotes and to what extent RNA degradation occurs co-translationally. Although cryo-EM data was collected on endogenous ribosome-Ski complexes isolated from wild-type *S. cerevisiae* (Schmidt et al., 2016), the structure of yeast Xrn1 bound to the ribosome was specifically obtained from *ski2*Δ cells, by trapping the complex on a reporter construct containing ribosome stalling sequences (Tesina et al., 2019). Thus, while Xrn1 was shown to co-sediment with polysomes (Tesina et al., 2019), it was unclear if direct interaction between the exonuclease and the ribosome would also occur on canonical mRNAs and under normal conditions. Furthermore, another study in *Drosophila* S2 cells suggests that the association between Xrn1 and ribosomes is not direct, as it was sensitive to treatment with micrococcal nuclease (Antic et al., 2015). However, ribosomal RNA hits from our Xrn1 and Skiv2l CRAC data clearly show that the two decay factors crosslink to the 40S subunit, indicating that they can directly interact with the ribosome. Moreover, the crosslinking sites on the small ribosomal subunit that we identified for Xrn1 and Skiv2l are in good agreement the cryo-EM structures from *S. cerevisiae*. Thus, we demonstrate that the direct association between the two decay factors and the ribosome is conserved in mammalian cells and more importantly, that it can occur under wild-type conditions.

When examining the global mRNA binding patterns of Xrn1 and Skiv2l, we noticed prominent differences in the distribution of these two factors along transcripts. Consistent with previous transcriptome-wide studies in yeast (Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013), our data indicate that Xrn1 is enriched in the 3′ UTR. Since the interaction between the ribosome and Xrn1 has been suggested to contribute to the processivity of the exonuclease (Tesina et al., 2019), it is possible that increased binding to the 3′ UTR reflects slower progression though this ribosome-free region, where the presence of secondary structures or bound proteins may negatively affect the processivity of Xrn1. Interestingly, Skiv2l binding along mRNAs displayed the opposite pattern. While the helicase was enriched at 5′ UTRs and along the coding sequence, it was almost completely excluded from the 3′ UTRs, closely mirroring the ribosome

distribution along a transcript. This binding pattern was rather surprising, as previous studies have shown that the Ski complex is required for efficient degradation of 3′ UTRs in yeast (Anderson and Parker, 1998). Further supporting this notion, transcriptome-wide data sets examining Ski2 binding along mRNAs in *S. cerevisiae*, clearly indicate that the helicase accumulates in the 3′ end of transcripts, including 3′ UTRs, which is at odds with our CRAC data (Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013). These differences can be reconciled by the recent discovery of a novel Ski-complex interacting factor, Ska1, which associates with a subpopulation of the Ski complex in yeast (Zhang et al., 2019). Although Ska1 has no enzymatic activity, it was shown to be required for Ski complex-mediated degradation of ribosome-poor regions including long 3′ UTRs and coding sequences with low ribosome occupancy (Zhang et al., 2019). As direct association with the ribosome is important for disrupting the autoinhibitory conformation of the Ski complex and for stimulating the catalytic activity of the helicase (Schmidt et al., 2016), it is not surprising that the complex may require similar assistance from other factors in the absence of nearby ribosomes. This model however lacks direct experimental evidence and it is equally likely that the sole function of Ska1 is to antagonize the interaction between the ribosome and the Ski complex. Consistent with this, Ski3 is displaced from polysomes upon overexpression of Ska1 (Zhang et al., 2019). Intriguingly, no Ska1 homologs are predicted to exist outside of other *Saccharomyces* species (Zhang et al., 2019), which may explain why the mammalian Skiv2l is exclusively associated with ribosome-occupied regions. Thus, our findings highlight a previously undescribed difference between the mammalian and the yeast 3′-5′ decay pathway.

The stark exclusion of Skiv2l form 3′UTRs suggests that the Ski complex is likely not involved in the degradation of full-length transcripts unlike Xrn1, which accumulates at the 3′ end of mRNAs. This would be consistent with a model where the 5′-3′ pathway also represents the main route for general mRNA turnover in mammalian cells, while the Ski complex and exosome are exclusively involved in more specialized decay triggered by translational events, including mRNA surveillance. In support of this, we found that compared to Xrn1, Skiv2l targets were characterized by higher ribosome occupancy and Skiv2l binding was diminished upon treatment with the translational inhibitors cycloheximide and harringtonine. These observations strongly suggest that the ribosome is a major determinant of Skiv2l recruitment in mammalian cells. On the other hand, Xrn1-mediated degradation in yeast was not strongly impacted by decoupling the exonuclease from the ribosome (Tesina et al., 2019), suggesting that the physical link with translation might be a byproduct, rather than a prerequisite for 5′-3′ decay.

We cannot fully exclude the possibility that the exosome degrades full-length transcripts including 3′ UTRs independently of its helicase co-factor. This seems unlikely given that all reported functions of the cytoplasmic exosome require the Ski complex (Anderson and Parker, 1998; Araki et al., 2001; Mitchell and

Tollervey, 2003; Orban and Izaurralde, 2005; van Hoof et al., 2002; van Hoof et al., 2000). Furthermore, in the absence of Ski2, the exosome could only trim the 3′ ends of transcripts by about 50 nucleotides from the poly(A) site (Tesina et al., 2019), suggesting that at least in yeast, the function of the Ski complex is required for efficient exonucleolytic activity of the exosome. Of note, our CRAC data indicate that the mammalian Skiv2l can in fact bind along very short 3′UTRs, possibly because this length still allows it to engage nearby ribosomes in the ORF. Consistent with this, ribosomes with a short 3′ RNA overhang of 10-20 and up to 50 nucleotides were shown to interact with the yeast Ski complex (Schmidt et al., 2016).

Nevertheless, it could be interesting to examine whether similar to yeast, all functions of the mammalian cytoplasmic exosome are coupled to its co-factor. This could be achieved by initial comparison of Skiv2l CRAC data sets versus Dis3l, the catalytic subunit of cytoplasmic exosome. If the two complexes function together, their substrates as well as mRNA binding patterns should largely overlap. Conversely, if the exosome was involved in degradation of full-length transcripts independently of Skiv2l, Dis3l would be expected to crosslink along 3′ UTRs. These experiments could be complemented by examining transcript half-lives in in *Skiv2l*$^{-/-}$ versus *Dis3l*$^{-/-}$ cells, where any changes in the stability of direct targets should largely correlate between the two conditions. Similarly, performing Skiv2l CRAC in the absence of Dis3l could reveal whether coupling to the exonucleolytic activity of the exosome is required for the observed distribution of Skiv2l along mRNAs. It should be noted however, that binding of the yeast Ski complex to the ribosome is proposed to be the first step in initiating 3′-5 decay, as it can occur independently of the exosome (Schmidt et al., 2016). Therefore, the initial recruitment of the Ski complex to substrates is unlikely to be affected in *Dis3l*$^{-/-}$ cells, but depending on whether the helicase can move along the mRNA in the absence of an exonucleolytic activity, its binding pattern along the transcript might change.

Unlike *S. cerevisiae*, *S. pombe* and higher eukaryotes possess an additional 3′-5′ cytoplasmic exonuclease, Dis3l2, which operates independently of the exosome. Although Dis3l2 has been implicated in general mRNA decay in human cells and *S. pombe* (Lubas et al., 2013; Malecki et al., 2013), other studies suggest that it primarily functions as a quality control pathway for cytoplasmic non-coding RNAs (Łabno et al., 2016b; Pirouz et al., 2016; Ustianenko et al., 2016). Human Dis3l2 was found to co-sediment with polysomes, suggesting it might also degrade its mRNA targets co-translationally (Lubas et al., 2013). Thus, it would be interesting to test if the mammalian Dis3l2 complements the Ski-exosome 3′-5′ pathway by degrading full-length transcripts including their 3′UTRs under normal conditions. For this purpose, a systematic profiling of direct Dis3l2 mRNA targets by CRAC and RNA-seq analysis of *Dis3l2*$^{-/-}$ cells could useful in examining any distinct and overlapping functions with the other cytoplasmic pathways.

## 2. Skiv2l binding reveals triggers of translation-coupled RNA decay

Consistent with its close links to the ribosome, we found that Skiv2l was enriched at sites that are known to trigger RNA degradation as a result of abnormal translation. For instance, uORFs known to induce NMD were occupied by Skiv2l (e.g. Ifrd1), and other endogenous NMD substrates (e.g. alternatively-spliced Dhx9) were bound by Skiv2l upstream of the premature stop codon, consistent with a role for the helicase in degrading the 5′ RNA fragment following an endonucleolytic cleavage. To examine if Skiv2l responds to other aberrant translation events, we combined our CRAC data with disome profiling, which we used as a proxy for prolonged translational stalls that are likely to trigger mRNA surveillance. We observed disome enrichment at regions encoding stretches of poly-lysine, -proline, -glutamate, -aspartate and -arginine which were previously reported to pause ribosomes (Ingolia et al., 2011). Skiv2l binding was elevated over the same regions, suggesting that these sequences are indeed more susceptible to surveillance. However, for some amino acid stretches, such as poly-prolines, the Skiv2l occupancy was much less pronounced than the disome peak, suggesting that certain pauses are resolved without triggering surveillance and decay of the message. This is consistent with the presence of specific factors such as eIF5A, which has been shown to alleviate translational stalls at poly-proline stretches by facilitating the reactivity of the amino acid during peptide bond formation (Gutierrez et al., 2013; Pelechano and Alepuz, 2017; Schuller et al., 2017). Of note, the 3′ ends of Skiv2l-bound RNA fragments at proline pairs were enriched in frame, indicating that the incorporation of this amino acid into the nascent peptide, and not the underlying mRNA sequence, is responsible for triggering surveillance.

Translational arrest on poly-basic (lysine and arginine) encoding sequences has been proposed to occur as a result of electrostatic interactions between the positively charged polypeptide and the negatively charged ribosome exit tunnel (Lu and Deutsch, 2008). However, the degree to which poly-basic residues impede translation varies greatly between organisms. While consecutive stretches of arginines are potent inducers of surveillance due to translational arrest in yeast (Brandman et al., 2012; Ikeuchi et al., 2019; Letzring et al., 2013), poly-arginines have little effect in mammalian cells. Instead, the main triggers of ribosome stalls in human cells appear to be poly-lysines specifically encoded by consecutive AAA codons (Garzia et al., 2017; Juszkiewicz et al., 2018; Lu and Deutsch, 2008; Sundaramoorthy et al., 2017). These findings suggest that elongation pausing at polybasic residues depends on the underlying mRNA sequence, rather than incorporation of the specific amino acid. This is consistent with previous work demonstrating that reporters with iterated AAA codons produced significantly less protein than those with equivalent number of AAGs, despite coding for the same poly-lysine peptide (Arthur et al., 2015; Koutmou et al., 2015). In line with these findings, we observed equivalent Skiv2l binding at lysine-rich versus lysine-poor A/G sequences, while Skiv2l and disomes were enriched at A-rich but not G-rich stretches. This suggest that adenine-rich

mRNA sequences are prone to translation-coupled surveillance involving Skiv2l, and that sequence, rather than the encoded amino acid, is key.

Since ubiquitination of stalled elongation complexes has been shown to trigger mRNA surveillance and protein quality control, it is tempting to speculate that Skiv2l might recognize such ubiquitination events. Of note, the arch domain of Skiv2l contacts the 40S ribosomal subunits uS3 (Rps3) and uS10 (Rps20) (Schmidt et al., 2016), both of which are substrates for ubiquitination (Garzia et al., 2017; Ikeuchi et al., 2019; Juszkiewicz and Hegde, 2017; Sundaramoorthy et al., 2017). However, no ubiquitin densities were detected in the cryo-EM structure of the ribosome bound by the Ski complex, suggesting that this modification is likely not essential for Skiv2l binding. Nevertheless, it is possible that ubiquitinated ribosomal subunits might be recognized by the recently discovered NGD/NSD endonuclease, Cue2, which has several ubiquitin binding domains (D'Orazio et al., 2019). The mammalian Cue2 ortholog, N4bp2, is also predicted to bind ubiquitin, and it would be interesting to test if N4bp2 is the endonuclease that acts upstream of the Ski complex to cleave mRNAs during aberrant translation and create 'access points' for Skiv2l within the ORF to initiate 3′-5 decay. This could be tested by performing Skiv2l CRAC in the absence of N4bp2. If this factor is indeed the primary endonuclease acting in response to stalled ribosomes, Skiv2l recruitment to mRNAs and especially at sites that trigger surveillance should be diminished.

## 3. Skiv2l determines the steady-state level of a subset of mRNAs

Although our data support a model where Xrn1 is the major exonuclease for general mRNA turnover, whereas the Ski complex specializes in translation-trigged decay and surveillance, we wondered whether there are certain transcripts which would be primarily degraded through the 3′-5′ pathway. Consistent with the idea that mRNAs can be targeted for decay both in the 5′-3′ and 3′-5′ direction, on many individual transcripts we detected similar Xrn1 and Skiv2l binding. However, a more global comparison of the mRNAs bound by Xrn1 and Skiv2l using PCA analysis revealed prominent and reproducible differences. This pattern could also be observed on many individual transcripts where binding of one decay factor seemed to dominate. This was particularly intriguing for the set of mRNAs highly bound by Skiv2l, but not Xrn1, as our CRAC data would suggest that they are primarily targeted by the Ski-exosome pathway. Consistent with this, many mRNAs highly bound by Skiv2l were also upregulated upon Skiv2l knock out, suggesting that the 3′-5′ pathway is important for determining their steady-state levels. This group includes the replication-dependent histones, which represent some of the most highly upregulated transcripts in the Skiv2l knock out cells.

Certain features of the replication-dependent histone mRNAs could make them optimal Ski complex substrates. For instance, their very short 3′ UTRs may allow Skiv2l to engage the 3′ end of the full-length transcript by enabling interactions with ribosomes from the nearby ORF. Indeed, we often observe Skiv2l CRAC signal along the whole length of the histone 3′UTR, including over the entire stem-loop structure. These observations suggest that 3′-5′ histone mRNA decay through the Ski complex may bypass the need for the 3′-5′ exonuclease Eri1, which was shown to be required for efficient degradation of the stem-loop, following uridylation (Hoefig et al., 2013). Consistent with this, some studies report that knock down of Eri1 has no discernible effect on histone mRNA stability (Mullen and Marzluff, 2008). In addition, it would be interesting to test if Upf1 is dispensable for Ski complex-mediated decay, since Skiv2l also has ATPase and helicase activities. Finally, it has been shown that active translation of histone mRNAs is required for their rapid degradation upon inhibition of DNA synthesis (Graves et al., 1987; Kaygun and Marzluff, 2005b). This feature is likely to further contribute to the affinity of Skiv2l for histone mRNAs, as our data indicate that ribosome occupancy is a prerequisite for Skiv2l binding.

Efficient histone mRNA degradation has been previously shown to require the activity of the exosome through its catalytic subunit Rrp6 (Mullen and Marzluff, 2008; Slevin et al., 2014). This is interesting given that Rrp6 primarily functions in the nucleus, although small amounts of it have also been detected in the cytoplasm (Burkard and Butler, 2000; Januszyk et al., 2011). Thus, it could be interesting to reevaluate the relative contribution of the exosome exonucleases, although it might not be surprising that multiple factors and pathways are involved in histone mRNA decay to confer robustness to the process.

## 4. Aven and Focadhesin interact with the Ski complex

In stark contrast to the variety of auxiliary proteins that associate with the nuclear exosome, the Ski complex is the only described co-factor that assists the exosome in the cytoplasm. In yeast, a novel Ski complex-interacting factor, Ska1, was recently shown to associate with a subset of Ski-exosome complexes to facilitate degradation of ribosome-devoid RNAs (Zhang et al., 2019). Although Ska1 is not conserved outside of the fungal species, its discovery raises the question of whether additional unidentified factors may function with the Ski complex in other organisms as well. Our work identified two proteins, Aven and Focadhesin, which interact with the Ski complex in mES cells, thus expanding the set of factors that coordinate cytoplasmic 3′-5′ RNA decay in mammalian cells.

Focadhesin is a large cytoplasmic protein initially identified as a focal adhesion component that functions as tumor suppressor in glioblastoma (Brockschmidt et al., 2012). Focadhesin was shown to interact and co-localize with the focal adhesion protein vinculin and its re-expression in glioblastoma cells led to decreased

proliferation and cell invasion both *in vitro* and in *vivo* (Brockschmidt et al., 2012). Focadhesin deletions have also been associated with breast and colorectal cancer (Weren et al., 2015). Further characterization of the molecular functions of Focadhesin as a tumor suppressor revealed that it interacts with tubulin and impacts microtubule dynamics during the cell cycle consistent with its localization to the centrosome (Brand et al., 2019). Focadhesin knock down also led to accelerated progression through the G2/M phase of the cell cycle (Brand et al., 2019).

Interestingly, a recent study showed that the plant Focadhesin homolog, Rst1, together with another protein of unknown function, Ripr, associate with the exosome and the Ski complex (Lange et al., 2019). More specifically, IP-MS analysis revealed that Rst1 mainly co-purifies with its binding partner Ripr, Ski7 and the core exosome, while the Ski complex components and Ski7 were most highly enriched in the Ripr co-IPs (Lange et al., 2019). This suggests that Rst1 and its binding partner Ripr bridge the interaction between the Ski7-exosme complex and the Ski2-Sk3-Ski8 proteins. Ripr is a 356 amino acid protein without any apparent functional domains or resemblance to other proteins (Lange et al., 2019). The phenotypes associated with loss of function of either Rst1, Ripr or the exosome were largely overlapping, suggesting that all components might act in the same pathway (Lange et al., 2019). Furthermore, similar to exosome and Ski complex mutants, loss of Rst1 and Ripr led to production of secondary siRNAs from certain exosome substrates such as endogenous mRNAs or RISC-cleavage products (Lange et al., 2019). Therefore, these observations strongly support a role for Rst1 and Ripr as co-factors of the cytoplasmic exosome and the Ski complex in plants (Lange et al., 2019).

Aven is a primarily cytoplasmic protein, which is expressed in a variety of tissues and cell lines (Chau et al., 2000). It was initially identified in a yeast two-hybrid screen as a protein that interacts with and stabilizes the pro-survival factor Bcl-xL (Chau et al., 2000). The same study also showed that Aven binds to Apaf1 (apoptotic protease activating factor 1) and prevents its oligomerization, which inhibits formation of the apoptosome, activation of Caspase-9 and subsequent cell death, thus establishing Aven as an apoptosis inhibitor (Chau et al., 2000). Subsequent studies showed that proteolytic cleavage of the N-terminus of Aven by Cathepsin-D is required to activate its pro-survival properties (Melzer et al., 2012). Consistent with its function as an apoptotic inhibitor, Aven overexpression has been linked to the development of acute leukemias and correlates with poor prognosis in childhood acute lymphoblastic leukemia (Choi et al., 2006; Eißmann et al., 2012). Aven was also implicated in cell cycle regulation in response to DNA damage. The addition of Aven mRNA to cycling *Xenopus* egg extracts caused activation of the ataxia telangiectasia-mutated (ATM) kinase and inhibited progression through the G2/M phase (Guo et al., 2008). On the other hand, loss of Aven in osteosarcoma cells induced G2 cell cycle arrest and reduced activation of ataxia telangiectasia-related (ATR) kinase in response to DNA damage (Baranski et al., 2015).

A more recent study of the molecular functions of Aven showed that it co-fractionates with polysomes and binds RNAs through its N-terminal RGG/RG box, which is a known RNA-interaction motif, found in various RNA binders (Thandapani et al., 2013; Thandapani et al., 2015). Arginine residues within RGG/RG sequences can be methylated by protein arginine methyltransferases (PRMTs) to modulate the properties of the motif (Thandapani et al., 2013). Similarly, Aven association with polysomes is dependent on methylation of its RGG/RG box by Prmt1 (Thandapani et al., 2013). It was also shown that Aven recognizes G-quadruplex RNAs (Thandapani et al., 2015), which are higher-order RNA structures formed by stacks of guanine tetrads that are stabilized by potassium ions (Kwok and Merrick, 2017). *In vivo*, Aven was bound to the G-quadruplex motif within the ORFs of two mRNAs encoding for the mixed lineage leukemia family proteins (Thandapani et al., 2015). This increased the polysomal association of the mRNAs and facilitated their translation (Thandapani et al., 2015). Recruitment of the helicase Dhx36 implicated in resolving G-quadruplex structures was suggested to mediate the positive effect of Aven on translation (Creacy et al., 2008; Lattmann et al., 2010; Thandapani et al., 2015; Tippana et al., 2019; Yangyuoru et al., 2018). Interestingly, Aven IP-MS data sets from this study indicate interactions with Focadhesin as well as all Ski complex components (Thandapani et al., 2015), which is in line with our findings, although we did not detect any association between Aven and Dhx36 or Prmt1. Our data indicate that Aven and Focadhesin are stably bound and interact more weakly or transiently with the Ski complex.

## 5. Aven and the Ski complex act in concert to oppose translational stalling

The interaction of Aven and Focadhesin with the Ski complex raised the question of whether these two factors are involved in cytoplasmic 3′-5′ RNA decay. Since Focadhesin was challenging to work with due to its large size and low abundance, we mainly focused on characterizing the interaction between Aven and the Ski complex. Global profiling of Aven substrates by CRAC revealed that Skiv2l and Aven share similar mRNA and ribosome binding sites, supporting a role for these proteins in the same pathway. Perhaps our most striking finding is the pronounced global increase of Skiv2l binding to RNAs upon Aven knock out. We observed similar changes in the absence of Focadhesin, suggesting that Aven and Focadhesin have overlapping functions. As our previous results indicate that the distribution of Skiv2l across transcripts is tightly coupled to ribosomes, we predicted that the changes in Skiv2l binding in the absence of Aven could presumably reflect altered ribosome occupancy. Indeed, mRNAs that gained Skiv2l also displayed an increase in disome occupancy, suggesting that ribosomes were more prone to stalling in the absence of Aven. Collectively, these observations are consistent with a model where Aven and Skiv2l serve complementary roles as part of a translation-coupled RNA surveillance pathway, where Aven opposes

prolonged translational stalls and Skiv2l assists in eliminating transcripts where these aberrant events accumulate (Figure 17).

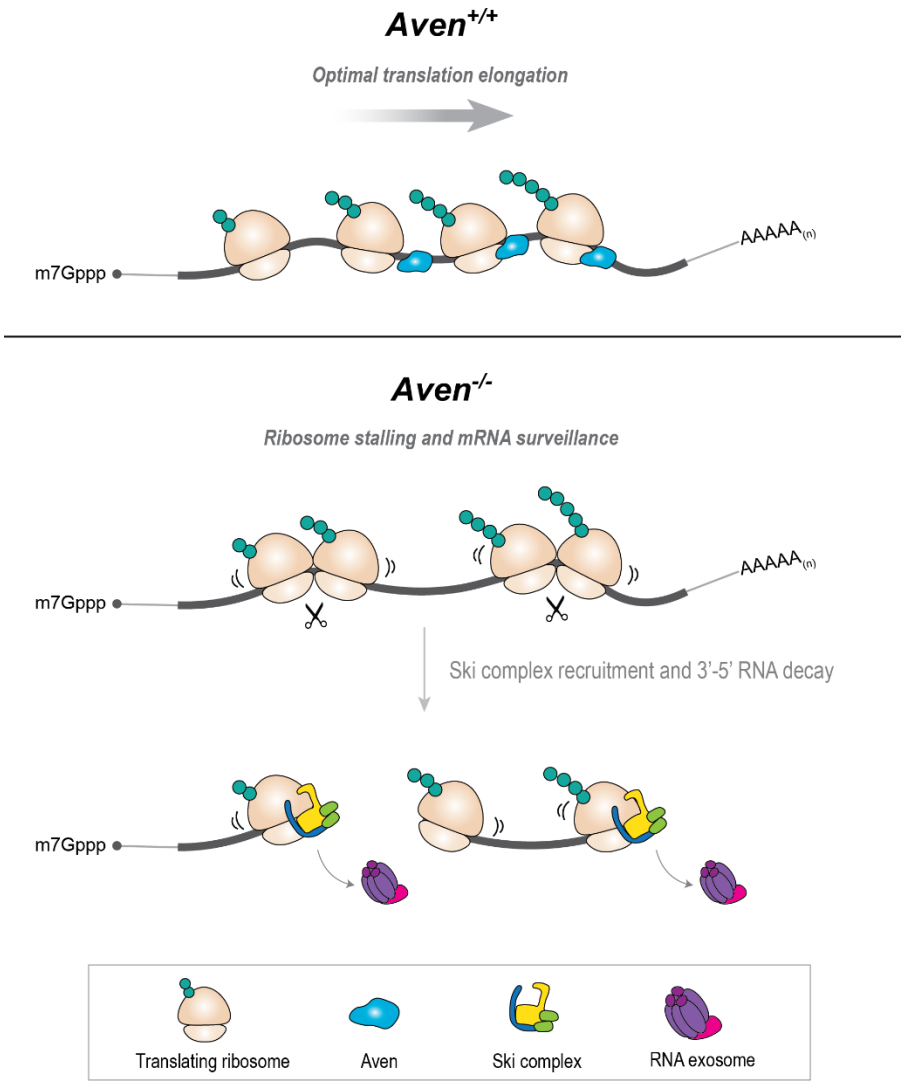



**Figure 17. Model for how the combined functions of Aven and the Ski complex could oppose translational stalling.**

In WT cells, Aven associates with polysomes by direct binding to the 40S ribosomal subunit and/or the mRNA. This facilitates ribosome translocation along difficult to translate sequences, including structured regions. In the absence of Aven, ribosomes are more susceptible to stalling which promotes the formation of collided ribosomes. The stalled mRNPs signal to the mRNA surveillance machinery which initiates an endonucleolytic cleavage of the message at the vicinity of the stalled ribosome. This event creates an 'access point' for Skiv2l which recognizes and binds the 80S ribosome presenting a short 3′ mRNA overhang. RNA degradation in the 3′-5′ direction is presumably initiated by subsequent recruitment of the RNA exosome to the Ski-ribosome complex.

Several lines of evidence further support this proposed model. Aven was enriched at RNA sequences predicted to have higher structural propensity, including G-quadruplexes and GC-rich sequences. Since structured RNA regions can negatively impact translation elongation, Aven could potentially facilitate ribosome translocation along the mRNA by helping to resolve such structural elements. This would be consistent with a reported role for Aven in promoting the optimal translation of mRNAs containing G-quadruplexes, through recruitment of the helicase Dhx36 (Thandapani et al., 2015). However, besides Skiv2l, we did not detect any other significantly enriched helicases in our Aven IPs. Nevertheless, it might be possible that the RGG/RG domain of Aven directly contributes to the remodeling of the structured region as previously suggested for other RGG/RG motifs (Loughlin et al., 2019; Meyer et al., 2019). Of note, in the Aven knock out, Skiv2l binding was elevated at predicted structural sequences, indicating that they are indeed more susceptible to surveillance. On a more global scale, elevated Skiv2l binding in the absence of Aven was paralleled by an increase in Skiv2l CRAC reads containing non-templated U-tails, which are thought to be added to RNA 3′ ends to help initiate decay. This further supports the notion that lack of Aven leads to increased surveillance and 3′-5′ RNA decay.

One prediction from our model is that the combined absence of Aven and Skiv2l should have an additive effect on transcript levels, as Skiv2l would not be available to clear stalled mRNPs forming due to lack of Aven. Consistent with this, we observed a more significant upregulation of Aven-Skiv2l substrates in the double knock compared to a single Skiv2l deletion. This trend was best illustrated by the histone mRNAs and small-ORF (smORF) containing RNAs originating from intergenic, upstream or antisense loci. A common feature of these two sets of transcripts is their high GC content. In addition, the smORF RNA sequences were predicted to form secondary structures. These characteristics might explain why histone mRNAs and smORF RNAs are particularly dependent on the Aven-Skiv2l pathway.

According to our model, Aven functions upstream of Skiv2l to oppose prolonged translational pausing, while Skiv2l acts at a later step to clear stalled mRNPs. Since this scenario would not necessarily require direct association between Aven and the Ski complex, it is perhaps somewhat puzzling that they interact, although they do not seem to form a stable complex. In light of these observations, it is important to consider alternative models that could also explain our data. For instance, since Aven and Skiv2l seem to occupy a similar site on the ribosome, near the mRNA entry channel, could it be possible that Aven competes with Skiv2l for binding to the 40S subunit? In this case, Aven could act to limit the association of the Ski complex with polysomes, by masking its binding site on the ribosome (Figure 18). This could also explain why Skiv2l binding to mRNAs is increased in the absence of Aven. However, if the sole purpose of Aven was to exclude Skiv2l from polysomes, it is unlikely that its loss would also affect monosome and disome

distribution. Nevertheless, this question could be addressed biochemically by testing Ski complex binding to the ribosome in the presence and absence of Aven.

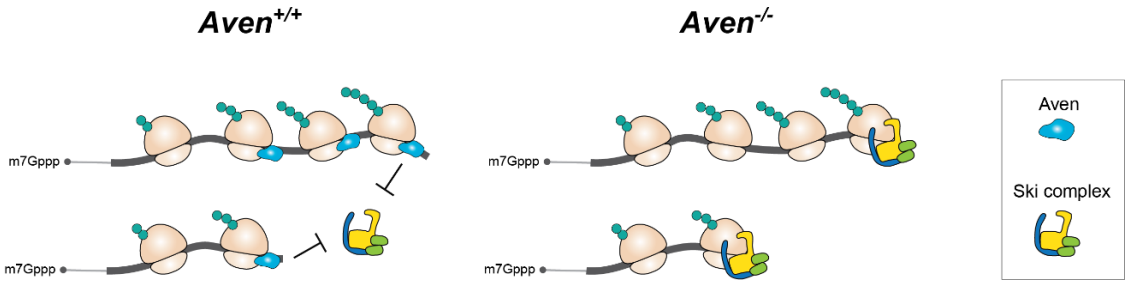



**Figure 18. Model for competition between Aven and the Ski complex**
Aven could potentially mask the Skiv2l binding site on the 40S ribosomal subunit to prevent or minimize the association of the Ski complex with polysomes. The absence of Aven could therefore expose the Skiv2l binding site on the 40S subunit and potentially facilitate recruitment of the Ski complex to the translating ribosome.

In an alternative scenario, Aven could modulate the enzymatic properties of the Ski complex. For instance, Aven might act to stabilize the closed conformation of the Ski complex and further limit its catalytic activity, even when the Ski complex is in contact with the ribosome (Figure 19). Conversely, in the absence of Aven, the autoinhibitory interactions within the Ski complex might be easier to remodel upon binding to the ribosome, leading to stabilization of Skiv2l on the mRNA and promoting its catalytic activity. This model would reconcile the interaction between Aven and the Ski complex as well as the increase in Skiv2l binding and 3′-5′ RNA decay in the absence of Aven, although it still does not explain the observed changes in ribosome occupancy. To examine if Aven directly affects the catalytic activity of Skiv2l, we could test the RNA binding, helicase and ATPase properties of the Ski complex in isolation and in combination with Aven.

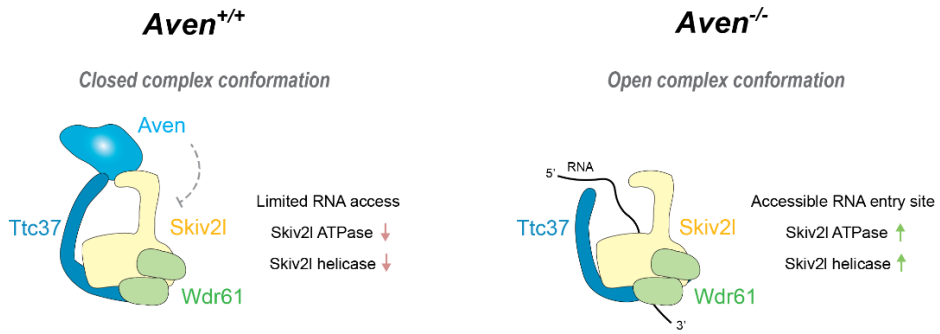



**Figure 19. Model for how Aven might affect the enzymatic activity of the Ski complex**
The interaction between Aven and the Ski complex could potentially stabilize the autoinhibitory conformation of the Ski complex and thus limit its RNA binding and ATPase and helicase activities. In this case, the absence of Aven could facilitate activation of the Ski complex and promote it catalytic activity.

Finally, Aven might affect the association between the Ski complex and the exosome, for instance by promoting or restricting exosome recruitment (Figure 20). The interaction between the Ski complex and the exosome is typically more transient and its detection in co-IP experiments often relies on prior crosslinking of the sample (Kalisiak et al., 2017). Therefore, to test if Aven controls the association of the two complexes, we could perform Skiv2l IP-MS experiments under crosslinking conditions, in the presence and absence of Aven. Given that the plant homolog of Focadhesin, Rst1, was recently shown to bridge interactions between the exosome and the Ski complex, it would also be important to test if the mammalian Focadhesin performs a similar function. Again, this could be addressed by Skiv2l IP-MS analysis in wild type versus Focadhesin knock out cells. Depending on the outcome of these experiments, it might also be important to test if the combined loss of both proteins affects the interaction between the Ski complex and the exosome, since our data suggest that Aven and Focadhesin have overlapping functions.

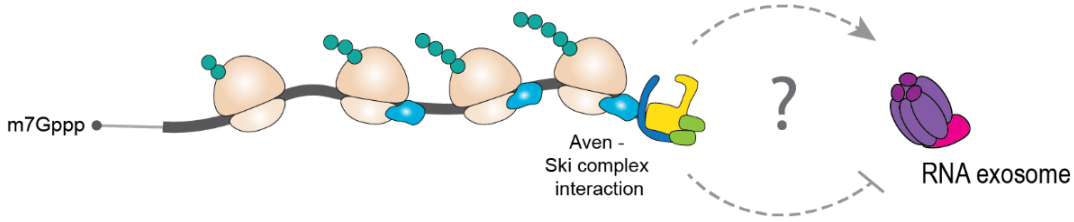



**Figure 20. Model for how Aven could influence the interaction between the Ski complex and the RNA exosome.** Binding of Aven (shown in light blue) to the Ski complex could potentially facilitate or limit the recruitment of the RNA exosome to either promote or oppose 3′-5′ RNA degradation of Skiv2l substrates.

In conclusion, we found that the two major cytoplasmic decay pathways in mammalian cells have physical links to translation. Our data suggest that general mRNA turnover mostly occurs via the Xrn1 5′-3′ pathway, but also reveal a subset of transcripts whose steady-state levels depend on the 3′-5′ pathway. We present evidence that the mammalian Ski complex helicase Skiv2l is exclusively recruited to ribosome-occupied regions, strongly suggesting that the Ski complex functions predominantly in translation-associated RNA degradation. We further identified Focadhesin and Aven as interaction partners of the Ski complex. Although additional biochemical and structural data would be necessary to elucidate the precise molecular functions of these interactions, our data are consistent with a model where Aven and the Ski complex have distinct but complementary roles in a translation coupled-surveillance pathway. While Aven seems to oppose ribosome stalling possibly by facilitating translation through structured regions, the Ski complex helps clear the transcript in case the stall cannot be resolved. Interestingly, the Aven-Skiv2l pathway acts on a wide range of substrates, including mRNAs, uORFs and most surprisingly on small ORF-containing RNAs derived from transcription of non-coding regions.

# Acknowledgements

First and foremost, I would like to thank my PhD advisor, Marc, for the opportunity to join the lab and work on these projects. I am extremely grateful for all his support throughout the years and I really appreciate his enthusiasm for science, his positive attitude and the fact that he is always encouraging and supportive. I am also thankful for all the freedom that Marc gives us to explore and work on different things. A huge thank you also goes to Alex Tuck with whom I shared my second project. Working with Alex has been just absolutely amazing!! I learned so much from him and I am immensely grateful for this fruitful collaboration. I also have to thank the entire Bühler lab for the amazing atmosphere and work environment. It was just such a pleasure to work with people who are so friendly, open, and always willing to help and share. I would also like to thank Matyas for letting us use his optimized protocols and for helping me getting started into the lab when I first joined. Also, special thanks to Yuki and Nathalie for all their hard work in keeping the lab in such a great order :) Furthermore, I am thankful to our amazing FMI facilities and especially Functional Genomics for sequencing our many CRAC and RNA-seq libraries. I also worked closely with Daniel and Vytautas from the Protein Analysis Facility, who contributed significantly to this work.

I would also like to thank Jeff Chao and Stefanie Jonas for being part of my PhD thesis committee and for all their help with the project. I am also grateful to Jessica and Varun from the Jeff's lab and to members of Stefanie's lab for their work on Aven and the Ski complex and for sharing ideas about the project. It was really fun to collaborate and meet so many people. This also includes David Gatfield, Bulak, Angelica and Violeta from the Gatfield lab at the University of Lausanne who helped a lot with the ribosome profiling part of the project.

# References


Adelman, K., Wei, W., Ardehali, M.B., Werner, J., Zhu, B., Reinberg, D., and Lis, J.T. (2006). Drosophila Paf1 modulates chromatin structure at actively transcribed genes. Mol Cell Biol *26*, 250-260.

Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. Nat Rev Genet *17*, 487-500.

Amrich, C.G., Davis, C.P., Rogal, W.P., Shirra, M.K., Heroux, A., Gardner, R.G., Arndt, K.M., and VanDemark, A.P. (2012). Cdc73 subunit of Paf1 complex contains C-terminal Ras-like domain that promotes association of Paf1 complex with chromatin. J Biol Chem *287*, 10863-10875.

Anderson, J.S., and Parker, R.P. (1998). The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. EMBO J *17*, 1497-1506.

Antic, S., Wolfinger, M.T., Skucha, A., Hosiner, S., and Dorner, S. (2015). General and MicroRNA-Mediated mRNA Degradation Occurs on Ribosome Complexes in Drosophila Cells. Mol Cell Biol *35*, 2309-2320.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. EMBO J *20*, 4684-4693.

Arthur, L.L., Pavlovic-Djuranovic, S., Koutmou, K.S., Green, R., Szczesny, P., and Djuranovic, S. (2015). Translational control by lysine-encoding A-rich sequences. Science Advances *1*, e1500154.

Bahrampour, S., and Thor, S. (2016). Ctr9, a Key Component of the Paf1 Complex, Affects Proliferation and Terminal Differentiation in the Developing Drosophila Nervous System. G3 (Bethesda) *6*, 3229-3239.

Baranski, Z., Booij, T.H., Cleton-Jansen, A.M., Price, L.S., van de Water, B., Bovee, J.V., Hogendoorn, P.C., and Danen, E.H. (2015). Aven-mediated checkpoint kinase control regulates proliferation and resistance to chemotherapy in conventional osteosarcoma. J Pathol *236*, 348-359.

Battle, D.J., and Doudna, J.A. (2001). The stem-loop binding protein forms a highly stable and specific complex with the 3' stem-loop of histone mRNAs. RNA *7*, 123-132.

Bayne, E.H., White, S.A., Kagansky, A., Bijos, D.A., Sanchez-Pulido, L., Hoe, K.L., Kim, D.U., Park, H.O., Ponting, C.P., Rappsilber, J.*, et al.* (2010). Stc1: a critical link between RNAi and chromatin modification required for heterochromatin integrity. Cell *140*, 666-677.

Becker, T., Armache, J.P., Jarasch, A., Anger, A.M., Villa, E., Sieber, H., Motaal, B.A., Mielke, T., Berninghausen, O., and Beckmann, R. (2011). Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. Nat Struct Mol Biol *18*, 715-720.

Beelman, C.A., Stevens, A., Caponigro, G., LaGrandeur, T.E., Hatfield, L., Fortner, D.M., and Parker, R. (1996). An essential component of the decapping enzyme required for normal rates of mRNA turnover. Nature *382*, 642-646.

Bernstein, E., Kim, S.Y., Carmell, M.A., Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A.A., Elledge, S.J., Anderson, K.V., and Hannon, G.J. (2003). Dicer is essential for mouse development. Nature Genetics *35*, 215-217.

Betz, J.L., Chang M Fau - Washburn, T.M., Washburn Tm Fau - Porter, S.E., Porter Se Fau - Mueller, C.L., Mueller Cl Fau - Jaehning, J.A., and Jaehning, J.A. (2002). Phenotypic analysis of Paf1/RNA polymerase II complex mutations reveals connections to cell cycle regulation, protein synthesis, and lipid and nucleic acid metabolism. Molecular Genetics and Genomics.

Bhandari, D., Guha, K., Bhaduri, N., and Saha, P. (2011). Ubiquitination of mRNA cycling sequence binding protein from Leishmania donovani (LdCSBP) modulates the RNA endonuclease activity of its Smr domain. FEBS Lett *585*, 809-813.

Boeck, R., Tarun, S., Jr., Rieger, M., Deardorff, J.A., Muller-Auer, S., and Sachs, A.B. (1996). The yeast Pan2 protein is required for poly(A)-binding protein-stimulated poly(A)-nuclease activity. J Biol Chem *271*, 432-438.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009). The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. Cell *139*, 547-559.

Brand, F., Forster, A., Christians, A., Bucher, M., Thome, C.M., Raab, M.S., Westphal, M., Pietsch, T., von Deimling, A., Reifenberger, G.*, et al.* (2019). FOCAD loss impacts microtubule assembly, G2/M progression and patient survival in astrocytic gliomas. Acta neuropathologica.

Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.W., Zhou, S., King, D., Shen, P.S., Weibezahn, J.*, et al.* (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. Cell *151*, 1042-1054.

Bresson, S., Tuck, A., Staneva, D., and Tollervey, D. (2017). Nuclear RNA Decay Pathways Aid Rapid Remodeling of Gene Expression in Yeast. Mol Cell *65*, 787-800 e785.

Brockschmidt, A., Trost, D., Peterziel, H., Zimmermann, K., Ehrler, M., Grassmann, H., Pfenning, P.N., Waha, A., Wohlleber, D., Brockschmidt, F.F.*, et al.* (2012). KIAA1797/FOCAD encodes a novel focal adhesion protein with tumour suppressor function in gliomas. Brain *135*, 1027-1041.

Buhler, M., Steiner, S., Mohn, F., Paillusson, A., and Muhlemann, O. (2006a). EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nat Struct Mol Biol *13*, 462-464.

Buhler, M., Verdel, A., and Moazed, D. (2006b). Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. Cell *125*, 873-886.

Burger, K., and Gullerova, M. (2018). Nuclear re-localization of Dicer in primary mouse embryonic fibroblast nuclei following DNA damage. PLoS Genet *14*, e1007151.

Burkard, K.T., and Butler, J.S. (2000). A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol Cell Biol *20*, 604-616.

Chakrabarti, S., Bonneau, F., Schussler, S., Eppinger, E., and Conti, E. (2014). Phospho-dependent and phospho-independent interactions of the helicase UPF1 with the NMD factors SMG5-SMG7 and SMG6. Nucleic Acids Res *42*, 9447-9460.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., and Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. Mol Cell *41*, 693-703.

Chamieh, H., Ballut, L., Bonneau, F., and Le Hir, H. (2008). NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. Nat Struct Mol Biol *15*, 85-93.

Chang, C.T., Bercovich, N., Loh, B., Jonas, S., and Izaurralde, E. (2014a). The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1. Nucleic Acids Res *42*, 5217-5233.

Chang, H., Lim, J., Ha, M., and Kim, V.N. (2014b). TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. Mol Cell *53*, 1044-1052.

Chang, J.H., Xiang, S., Xiang, K., Manley, J.L., and Tong, L. (2011). Structural and biochemical studies of the 5'-->3' exoribonuclease Xrn1. Nat Struct Mol Biol *18*, 270-276.

Chau, B.N., Cheng, E.H., Kerr, D.A., and Hardwick, J.M. (2000). Aven, a novel inhibitor of caspase activation, binds Bcl-xL and Apaf-1. Mol Cell *6*, 31-40.

Chen, C.Y., Gherzi, R., Ong, S.E., Chan, E.L., Raijmakers, R., Pruijn, G.J., Stoecklin, G., Moroni, C., Mann, M., and Karin, M. (2001). AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. Cell *107*, 451-464.

Chen, C.Y., and Shyu, A.B. (2011). Mechanisms of deadenylation-dependent decay. WIRES RNA *2*, 167-183.

Chen, F.X., Woodfin, A.R., Gardini, A., Rickels, R.A., Marshall, S.A., Smith, E.R., Shiekhattar, R., and Shilatifard, A. (2015). PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. Cell *162*, 1003-1015.

Chen, L., Muhlrad, D., Hauryliuk, V., Cheng, Z., Lim, M.K., Shyp, V., Parker, R., and Song, H. (2010a). Structure of the Dom34-Hbs1 complex and implications for no-go decay. Nat Struct Mol Biol *17*, 1233-1240.

Chen, L.L., Yang, L., and Carmichael, G.G. (2010b). Molecular basis for an attenuated cytoplasmic dsRNA response in human embryonic stem cells. Cell Cycle *9*, 3552-3564.

Chlebowski, A., Lubas, M., Jensen, T.H., and Dziembowski, A. (2013). RNA decay machines: the exosome. Biochim Biophys Acta *1829*, 552-560.

Cho, H., Han, S., Choe, J., Park, S.G., Choi, S.S., and Kim, Y.K. (2013). SMG5-PNRC2 is functionally dominant compared with SMG5-SMG7 in mammalian nonsense-mediated mRNA decay. Nucleic Acids Res *41*, 1319-1328.

Cho, H., Kim, K.M., and Kim, Y.K. (2009). Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. Mol Cell *33*, 75-86.

Choi, J., Hwang, Y.K., Sung, K.W., Kim, D.H., Yoo, K.H., Jung, H.L., and Koo, H.H. (2006). Aven overexpression: association with poor prognosis in childhood acute lymphoblastic leukemia. Leukemia research *30*, 1019-1025.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA (New York, NY) *13*, 998-1016.

Chowdhury, A., and Tharun, S. (2008). lsm1 mutations impairing the ability of the Lsm1p-7p-Pat1p complex to preferentially bind to oligoadenylated RNA affect mRNA decay in vivo. RNA *14*, 2149-2158.

Chu, X., Qin X Fau - Xu, H., Xu H Fau - Li, L., Li L Fau - Wang, Z., Wang Z Fau - Li, F., Li F Fau - Xie, X., Xie X Fau - Zhou, H., Zhou H Fau - Shen, Y., Shen Y Fau - Long, J., and Long, J. (2013). Structural insights into Paf1 complex assembly and histone binding. Nucleic Acids Res.

Chu, Y., Simic, R., Warner, M.H., Arndt, K.M., and Prelich, G. (2007). Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. EMBO J *26*, 4646-4656.

Collart, M.A., and Weiss, B. (2019). Ribosome pausing, a dangerous necessity for co-translational events. Nucleic Acids Res.

Colmenares, S.U., Buker, S.M., Buhler, M., Dlakic, M., and Moazed, D. (2007). Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. Mol Cell *27*, 449-461.

Creacy, S.D., Routh, E.D., Iwamoto, F., Nagamine, Y., Akman, S.A., and Vaughn, J.P. (2008). G4 resolvase 1 binds both DNA and RNA tetramolecular quadruplex with high affinity and is the major source of tetramolecular quadruplex G4-DNA and G4-RNA resolving activity in HeLa cell lysates. J Biol Chem *283*, 34626-34634.

D'Orazio, K.N., Wu, C.C.-C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. eLife *8*, e49117.

De Almeida, C., Scheer, H., Zuber, H., and Gagliardi, D. (2018). RNA uridylation: a key posttranscriptional modification shaping the coding and noncoding transcriptome. WIRES RNA *9*.

de Farias, S.T., Dos Santos Junior, A.P., Rego, T.G., and Jose, M.V. (2017). Origin and Evolution of RNA-Dependent RNA Polymerase. Front Genet *8*, 125.

de Jong, R.N., Truffault, V., Diercks, T., Ab, E., Daniels, M.A., Kaptein, R., and Folkers, G.E. (2008). Structure and DNA binding of the human Rtf1 Plus3 domain. Structure *16*, 149-159.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harbor perspectives in biology *4*, a013706.

Doma, M.K., and Parker, R. (2006). Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature *440*, 561-564.

Dominski, Z., Yang, X.C., Kaygun, H., Dadlez, M., and Marzluff, W.F. (2003). A 3' exonuclease that specifically interacts with the 3' end of histone mRNA. Mol Cell *12*, 295-305.

Doyle, M., Badertscher, L., Jaskiewicz, L., Guttinger, S., Jurado, S., Hugenschmidt, T., Kutay, U., and Filipowicz, W. (2013). The double-stranded RNA binding domain of human Dicer functions as a nuclear localization signal. RNA *19*, 1238-1252.

Dziembowski, A., Lorentzen, E., Conti, E., and Seraphin, B. (2007). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat Struct Mol Biol *14*, 15-22.

Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., and Jensen, T.H. (2009). SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. Nat Struct Mol Biol *16*, 49-55.

Eberle, A.B., Stalder, L., Mathys, H., Orozco, R.Z., and Muhlemann, O. (2008). Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS biology *6*, e92.

Eckmann, C.R., Rammelt, C., and Wahle, E. (2011). Control of poly(A) tail length. WIRES RNA *2*, 348-361.

Eißmann, M., Melzer, I.M., Fernández, S.B.M., Michel, G., Hrabě de Angelis, M., Hoefler, G., Finkenwirth, P., Jauch, A., Schoell, B., Grez, M.*, et al.* (2012). Overexpression of the anti-apoptotic protein AVEN contributes to increased malignancy in hematopoietic neoplasms. Oncogene *32*, 2586-2591.

Fabre, A., and Badens, C. (2014). Human Mendelian diseases related to abnormalities of the RNA exosome or its cofactors. Intractable Rare Dis Res *3*, 8-11.

Fabre, A., Charroux, B., Martinez-Vinson, C., Roquelaure, B., Odul, E., Sayar, E., Smith, H., Colomb, V., Andre, N., Hugot, J.-P.*, et al.* (2012). SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome. Am J Hum Genet *90*, 689-692.

Fasken, M.B., Leung, S.W., Banerjee, A., Kodani, M.O., Chavez, R., Bowman, E.A., Purohit, M.K., Rubinson, M.E., Rubinson, E.H., and Corbett, A.H. (2011). Air1 zinc knuckles 4 and 5 and a conserved IWRXY motif are critical for the function and integrity of the Trf4/5-Air1/2-Mtr4 polyadenylation (TRAMP) RNA quality control complex. J Biol Chem *286*, 37429-37445.

Fierz, B., Chatterjee, C., McGinty, R.K., Bar-Dagan, M., Raleigh, D.P., and Muir, T.W. (2011). Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. Nat Chem Biol *7*, 113-119.

Fiorini, F., Boudvillain, M., and Le Hir, H. (2013). Tight intramolecular regulation of the human Upf1 helicase by its N- and C-terminal domains. Nucleic Acids Res *41*, 2404-2415.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Flemr, M., and Buhler, M. (2015). Single-Step Generation of Conditional Knockout Mouse Embryonic Stem Cells. Cell Rep *12*, 709-716.

Flemr, M., Malik, R., Franke, V., Nejepinska, J., Sedlacek, R., Vlahovicek, K., and Svoboda, P. (2013). A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes. Cell *155*, 807-816.

Flury, V., Georgescu, P.R., Iesmantavicius, V., Shimada, Y., Kuzdere, T., Braun, S., and Buhler, M. (2017). The Histone Acetyltransferase Mst2 Protects Active Chromatin from Epigenetic Silencing by Acetylating the Ubiquitin Ligase Brl1. Mol Cell *67*, 294-307 e299.

Franks, T.M., Singh, G., and Lykke-Andersen, J. (2010). Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay. Cell *143*, 938-950.

Frischmeyer, P.A., van Hoof, A., O'Donnell, K., Guerrerio, A.L., Parker, R., and Dietz, H.C. (2002). An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. Science *295*, 2258-2261.

Fuchs, S.M., Kizer, K.O., Braberg, H., Krogan, N.J., and Strahl, B.D. (2012). RNA polymerase II carboxyl-terminal domain phosphorylation regulates protein stability of the Set2 methyltransferase and histone H3 di- and trimethylation at lysine 36. J Biol Chem *287*, 3249-3256.

Fukui, K., and Kuramitsu, S. (2011). Structure and Function of the Small MutS-Related Domain. Molecular biology international *2011*, 691735.

Gagnon, K.T., Li, L., Chu, Y., Janowski, B.A., and Corey, D.R. (2014). RNAi factors are present and active in human cell nuclei. Cell Rep *6*, 211-221.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat Rev Mol Cell Bio *8*, 113-126.

Garzia, A., Jafarnejad, S.M., Meyer, C., Chapat, C., Gogakos, T., Morozov, P., Amiri, M., Shapiro, M., Molina, H., Tuschl, T.*, et al.* (2017). The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. Nature communications *8*, 16056.

Gerbasi, V.R., Weaver, C.M., Hill, S., Friedman, D.B., and Link, A.J. (2004). Yeast Asc1p and mammalian RACK1 are functionally orthologous core 40S ribosomal proteins that repress gene expression. Mol Cell Biol *24*, 8276-8287.

Glover, M.L., Burroughs, A.M., Egelhofer, T.A., Pule, M.N., Aravind, L., and Arribere, J.A. (2019). NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. bioRxiv, 674358.

Graille, M., Chaillet, M., and van Tilbeurgh, H. (2008). Structure of yeast Dom34: a protein related to translation termination factor Erf1 and involved in No-Go decay. J Biol Chem *283*, 7145-7154.

Graves, R.A., Pandey, N.B., Chodchoy, N., and Marzluff, W.F. (1987). Translation is required for regulation of histone mRNA degradation. Cell *48*, 615-626.

Grewal, S.I. (2010). RNAi-dependent formation of heterochromatin and its diverse functions. Curr Opin Genet Dev *20*, 134-141.

Grima, D.P., Sullivan, M., Zabolotskaya, M.V., Browne, C., Seago, J., Wan, K.C., Okada, Y., and Newbury, S.F. (2008). The 5'-3' exoribonuclease pacman is required for epithelial sheet sealing in Drosophila and genetically interacts with the phosphatase puckered. Biology of the cell *100*, 687-701.

Grudzien-Nogalska, E., Jiao, X., Song, M.G., Hart, R.P., and Kiledjian, M. (2016). Nudt3 is an mRNA decapping enzyme that modulates cell migration. RNA *22*, 773-781.

Grudzien-Nogalska, E., and Kiledjian, M. (2017). New insights into decapping enzymes and selective mRNA decay. WIRES RNA *8*.

Guo, J.Y., Yamada, A., Kajino, T., Wu, J.Q., Tang, W., Freel, C.D., Feng, J., Chau, B.N., Wang, M.Z., Margolis, S.S.*, et al.* (2008). Aven-dependent activation of ATM following DNA damage. Curr Biol *18*, 933-942.

Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., Kim, J.R., Saini, P., Buskirk, A.R., and Dever, T.E. (2013). eIF5A promotes translation of polyproline motifs. Mol Cell *51*, 35-45.

Guydosh, N.R., and Green, R. (2014). Dom34 rescues ribosomes in 3' untranslated regions. Cell *156*, 950-962.

Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. RNA *23*, 749-761.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millan-Zambrano, G., Barkai, O., Chavez, S., Perez-Ortin, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000-1011.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The yeast ski complex: crystal structure and RNA channeling to the exosome complex. Cell *154*, 814-826.

Halbach, F., Rode, M., and Conti, E. (2012). The crystal structure of S. cerevisiae Ski2, a DExH helicase associated with the cytoplasmic functions of the exosome. RNA *18*, 124-134.

Harris, M.E., Bohni, R., Schneiderman, M.H., Ramamurthy, L., Schumperli, D., and Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol Cell Biol *11*, 2416-2424.

Hartley, J.L., Zachos, N.C., Dawood, B., Donowitz, M., Forman, J., Pollitt, R.J., Morgan, N.V., Tee, L., Gissen, P., and Kahr, W.H. (2010). Mutations in TTC37 cause trichohepatoenteric syndrome (phenotypic diarrhea of infancy). Gastroenterology *138*, 2388-2398.

Hayashi, A., Ishida, M., Kawaguchi, R., Urano, T., Murakami, Y., and Nakayama, J. (2012). Heterochromatin protein 1 homologue Swi6 acts in concert with Ers1 to regulate RNAi-directed heterochromatin assembly. Proc Natl Acad Sci USA *109*, 6159-6164.

Hoefig, K.P., and Heissmeyer, V. (2014). Degradation of oligouridylated histone mRNAs: see UUUUU and goodbye. WIRES RNA *5*, 577-589.

Hoefig, K.P., Rath, N., Heinz, G.A., Wolf, C., Dameris, J., Schepers, A., Kremmer, E., Ansel, K.M., and Heissmeyer, V. (2013). Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. Nat Struct Mol Biol *20*, 73-81.

Hogg, J.R., and Goff, S.P. (2010). Upf1 senses 3'UTR length to potentiate mRNA decay. Cell *143*, 379-389.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat Rev Genet *16*, 71-84.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225-229.

Huntzinger, E., Kashima, I., Fauser, M., Sauliere, J., and Izaurralde, E. (2008). SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. RNA *14*, 2609-2617.

Hurt, J.A., Robertson, A.D., and Burge, C.B. (2013). Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay. Genome Res *23*, 1636-1650.

Ibrahim, F., Maragkakis, M., Alexiou, P., and Mourelatos, Z. (2018). Ribothrypsis, a novel process of canonical mRNA decay, mediates ribosome-phased mRNA endonucleolysis. Nat Struct Mol Biol *25*, 302-310.

Iida, T., Nakayama, J., and Moazed, D. (2008). siRNA-mediated heterochromatin establishment requires HP1 and is associated with antisense transcription. Mol Cell *31*, 178-189.

Ikeuchi, K., and Inada, T. (2016). Ribosome-associated Asc1/RACK1 is required for endonucleolytic cleavage induced by stalled ribosome at the 3' end of nonstop mRNA. Scientific reports *6*, 28234.

Ikeuchi, K., Tesina, P., Matsuo, Y., Sugiyama, T., Cheng, J., Saeki, Y., Tanaka, K., Becker, T., Beckmann, R., and Inada, T. (2019). Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. EMBO J *38*, e100276.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Jalkanen, A.L., Coleman, S.J., and Wilusz, J. (2014). Determinants and implications of mRNA poly(A) tail size--does this protein make my tail look big? Semin Cell Dev Biol *34*, 24-32.

Januszyk, K., Liu, Q., and Lima, C.D. (2011). Activities of human RRP6 and structure of the human RRP6 catalytic domain. RNA *17*, 1566-1577.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5'-end capping. Nature *467*, 608-611.

Jinek, M., Coyle, S.M., and Doudna, J.A. (2011). Coupled 5' nucleotide recognition and processivity in Xrn1-mediated mRNA decay. Mol Cell *41*, 600-608.

Joazeiro, C.A.P. (2019). Mechanisms and functions of ribosome-associated protein quality control. Nat Rev Mol Cell Bio *20*, 368-383.

Jones, C.I., Zabolotskaya, M.V., and Newbury, S.F. (2012). The 5' --> 3' exoribonuclease XRN1/Pacman and its functions in cellular processes and development. WIRES RNA *3*, 455-468.

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 Is a Quality Control Sensor of Collided Ribosomes. Mol Cell *72*, 469-481 e467.

Juszkiewicz, S., and Hegde, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. Mol Cell *65*, 743-750.e744.

Kalisiak, K., Kuliński, T.M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K., and Dziembowski, A. (2017). A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. Nucleic Acids Res *45*, 2068-2080.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev *19*, 489-501.

Kaygun, H., and Marzluff, W.F. (2005a). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. Nat Struct Mol Biol *12*, 794-800.

Kaygun, H., and Marzluff, W.F. (2005b). Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. Mol Cell Biol *25*, 6879-6888.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat Rev Mol Cell Bio *17*, 227-239.

Kim, D.H., Villeneuve, L.M., Morris, K.V., and Rossi, J.J. (2006). Argonaute-1 directs siRNA-mediated transcriptional gene silencing in human cells. Nat Struct Mol Biol *13*, 793-797.

Kim, J., Guermah, M., and Roeder, R.G. (2010). The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. Cell *140*, 491-503.

Kim, Y.K., and Maquat, L.E. (2019). UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. RNA *25*, 407-422.

Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., and Green, R. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. eLife *4*.

Kowalik, K.M., Shimada, Y., Flury, V., Stadler, M.B., Batki, J., and Buhler, M. (2015). The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. Nature *520*, 248-252.

Kowalinski, E., Kogel, A., Ebert, J., Reichelt, P., Stegmann, E., Habermann, B., and Conti, E. (2016). Structure of a Cytoplasmic 11-Subunit RNA Exosome Complex. Mol Cell *63*, 125-134.

Kowalinski, E., Schuller, A., Green, R., and Conti, E. (2015). Saccharomyces cerevisiae Ski7 Is a GTP-Binding Protein Adopting the Characteristic Conformation of Active Translational GTPases. Structure *23*, 1336-1343.

Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S., and Greenblatt, J.F. (2002). RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol *22*, 6979-6992.

Kuroha, K., Akamatsu, M., Dimitrova, L., Ito, T., Kato, Y., Shirahige, K., and Inada, T. (2010). Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. EMBO Rep *11*, 956-961.

Kurosaki, T., Li, W., Hoque, M., Popp, M.W., Ermolenko, D.N., Tian, B., and Maquat, L.E. (2014). A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. Genes Dev *28*, 1900-1916.

Kurosaki, T., and Maquat, L.E. (2013). Rules that govern UPF1 binding to mRNA 3' UTRs. Proc Natl Acad Sci USA *110*, 3357-3362.

Kwok, C.K., and Merrick, C.J. (2017). G-Quadruplexes: Prediction, Characterization, and Biological Application. Trends in biotechnology *35*, 997-1013.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016a). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta *1863*, 3125-3147.

Łabno, A., Warkocki, Z., Kuliński, T., Krawczyk, P.S., Bijata, K., Tomecki, R., and Dziembowski, A. (2016b). Perlman syndrome nuclease DIS3L2 controls cytoplasmic non-coding RNAs and provides surveillance pathway for maturing snRNAs. Nucleic Acids Res *44*, 10437-10453.

Lackey, P.E., Welch, J.D., and Marzluff, W.F. (2016). TUT7 catalyzes the uridylation of the 3' end for rapid degradation of histone mRNA. RNA *22*, 1673-1688.

Lai, T., Cho, H., Liu, Z., Bowler, M.W., Piao, S., Parker, R., Kim, Y.K., and Song, H. (2012). Structural basis of the PNRC2-mediated link between mrna surveillance and decapping. Structure *20*, 2025-2037.

Lange, H., Ndecky, S.Y.A., Gomez-Diaz, C., Pflieger, D., Butel, N., Zumsteg, J., Kuhn, L., Piermaria, C., Chicher, J., Christie, M.*, et al.* (2019). RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in Arabidopsis. bioRxiv, 617894.

Lattmann, S., Giri, B., Vaughn, J.P., Akman, S.A., and Nagamine, Y. (2010). Role of the amino terminal RHAU-specific motif in the recognition and resolution of guanine quadruplex-RNA by the DEAH-box RNA helicase RHAU. Nucleic Acids Res *38*, 6219-6233.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. Embo j *19*, 6860-6869.

LeBlanc, J.J., and Beemon, K.L. (2004). Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging. Journal of virology *78*, 5139-5146.

Lebreton, A., Tomecki, R., Dziembowski, A., and Seraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993-996.

Lee, H.H., Kim, Y.S., Kim, K.H., Heo, I., Kim, S.K., Kim, O., Kim, H.K., Yoon, J.Y., Kim, H.S., Kim, D.J.*, et al.* (2007). Structural and functional insights into Dom34, a key component of no-go mRNA decay. Mol Cell *27*, 938-950.

Letzring, D.P., Wolf, A.S., Brule, C.E., and Grayhack, E.J. (2013). Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. RNA *19*, 1208-1217.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell *159*, 1365-1376.

Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell *127*, 1223-1237.

Liu, X., Kraus, W.L., and Bai, X. (2015). Ready, pause, go: regulation of RNA polymerase II pausing and release by cellular signaling pathways. Trends Biochem Sci *40*, 516-525.

Loh, B., Jonas, S., and Izaurralde, E. (2013). The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2. Genes Dev *27*, 2125-2138.

Loughlin, F.E., Lukavsky, P.J., Kazeeva, T., Reber, S., Hock, E.-M., Colombo, M., Von Schroetter, C., Pauli, P., Cléry, A., Mühlemann, O.*, et al.* (2019). The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. Mol Cell *73*, 490-504.e496.

Lu, J., and Deutsch, C. (2008). Electrostatics in the ribosomal tunnel modulate chain elongation rates. J Mol Biol *384*, 73-86.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol Cell *43*, 624-637.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. EMBO J *32*, 1855-1868.

Lykke-Andersen, S., Chen, Y., Ardal, B.R., Lilje, B., Waage, J., Sandelin, A., and Jensen, T.H. (2014). Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. Genes Dev *28*, 2498-2517.

Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Bio *16*, 665-677.

Lyons, S.M., Ricciardi, A.S., Guo, A.Y., Kambach, C., and Marzluff, W.F. (2014). The C-terminal extension of Lsm4 interacts directly with the 3' end of the histone mRNP and is required for efficient histone mRNA degradation. RNA (New York, NY) *20*, 88-102.

Ma, E., MacRae, I.J., Kirsch, J.F., and Doudna, J.A. (2008). Autoinhibition of human dicer by its internal helicase domain. J Mol Biol *380*, 237-243.

Makino, D.L., Schuch, B., Stegmann, E., Baumgartner, M., Basquin, C., and Conti, E. (2015). RNA degradation paths in a 12-subunit nuclear exosome complex. Nature *524*, 54-58.

Malecki, M., Viegas, S.C., Carneiro, T., Golik, P., Dressaire, C., Ferreira, M.G., and Arraiano, C.M. (2013). The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. Embo j *32*, 1842-1854.

Mangus, D.A., and Jacobson, A. (1999). Linking mRNA turnover and translation: assessing the polyribosomal association of mRNA decay factors and degradative intermediates. Methods *17*, 28-37.

Marshall, A.N., Montealegre, M.C., Jimenez-Lopez, C., Lorenz, M.C., and van Hoof, A. (2013). Alternative splicing and subfunctionalization generates functional diversity in fungal proteomes. PLoS Genet *9*, e1003376.

Marzluff, W.F., and Koreski, K.P. (2017). Birth and Death of Histone mRNAs. Trends in genetics : TIG *33*, 745-759.

Matsuo, Y., Ikeuchi, K., Saeki, Y., Iwasaki, S., Schmidt, C., Udagawa, T., Sato, F., Tsuchiya, H., Becker, T., Tanaka, K.*, et al.* (2017). Ubiquitination of stalled ribosome triggers ribosome-associated quality control. Nature communications *8*, 159.

McLennan, A.G. (2006). The Nudix hydrolase superfamily. Cell Mol Life Sci *63*, 123-143.

Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell *44*, 43-52.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Melzer, I.M., Fernandez, S.B., Bosser, S., Lohrig, K., Lewandrowski, U., Wolters, D., Kehrloesser, S., Brezniceanu, M.L., Theos, A.C., Irusta, P.M.*, et al.* (2012). The Apaf-1-binding protein Aven is cleaved by Cathepsin D to unleash its anti-apoptotic potential. Cell Death Differ *19*, 1435-1445.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, Jens S., Sandelin, A.*, et al.* (2016). Identification of a Nuclear Exosome Decay Pathway for Processed Transcripts. Mol Cell *64*, 520-533.

Meyer, A., Golbik, R.P., Sanger, L., Schmidt, T., Behrens, S.E., and Friedrich, S. (2019). The RGG/RG motif of AUF1 isoform p45 is a key modulator of the protein's RNA chaperone and RNA annealing activities. RNA Biol *16*, 960-971.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases. Cell *91*, 457-466.

Mitchell, P., and Tollervey, D. (2003). An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation. Mol Cell *11*, 1405-1413.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell *30*, 755-766.

Morris, K.V., Chan, S.W., Jacobsen, S.E., and Looney, D.J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. Science *305*, 1289-1292.

Morton, D.J., Kuiper, E.G., Jones, S.K., Leung, S.W., Corbett, A.H., and Fasken, M.B. (2018). The RNA exosome and RNA exosome-linked disease. RNA *24*, 127-142.

Mosimann, C., Hausmann, G., and Basler, K. (2006). Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with beta-catenin/Armadillo. Cell *125*, 327-341.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Much, C., Auchynnikava, T., Pavlinic, D., Buness, A., Rappsilber, J., Benes, V., Allshire, R., and O'Carroll, D. (2016). Endogenous Mouse Dicer Is an Exclusively Cytoplasmic Protein. PLoS Genet *12*, e1006095.

Mueller, C.L., Porter, S.E., Hoffman, M.G., and Jaehning, J.A. (2004). The Paf1 complex has functions independent of actively transcribing RNA polymerase II. Mol Cell *14*, 447-456.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev *22*, 50-65.

Nagaike, T., Logan, C., Hotta, I., Rozenblatt-Rosen, O., Meyerson, M., and Manley, J.L. (2011). Transcriptional activators enhance polyadenylation of mRNA precursors. Mol Cell *41*, 409-418.

Newbury, S., and Woollard, A. (2004). The 5'-3' exoribonuclease xrn-1 is essential for ventral epithelial enclosure during C. elegans embryogenesis. RNA *10*, 59-65.

Ng, H.H., Dole, S., and Struhl, K. (2003). The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. J Biol Chem *278*, 33625-33628.

Nielsen, M.H., Flygaard, R.K., and Jenner, L.B. (2017). Structural analysis of ribosomal RACK1 and its role in translational control. Cellular signalling *35*, 272-281.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol Cell *39*, 773-783.

Nordick, K., Hoffman, M.G., Betz, J.L., and Jaehning, J.A. (2008). Direct Interactions between the Paf1 Complex and a Cleavage and Polyadenylation Factor Are Revealed by Dissociation of Paf1 from RNA Polymerase II. Eukaryotic Cell *7*, 1158.

Ohnishi, T., Yamashita, A., Kashima, I., Schell, T., Anders, K.R., Grimson, A., Hachiya, T., Hentze, M.W., Anderson, P., and Ohno, S. (2003). Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. Mol Cell *12*, 1187-1200.

Ohrt, T., Muetze, J., Svoboda, P., and Schwille, P. (2012). Intracellular localization and routing of miRNA and RNAi pathway components. Curr Top Med Chem *12*, 79-88.

Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., Izumi, N., Hirahara, F., and Ohno, S. (2012). N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. Nucleic Acids Res *40*, 1251-1266.

Orban, T.I., and Izaurralde, E. (2005). Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA (New York, NY) *11*, 459-469.

Ostapcuk, V., Mohn, F., Carl, S.H., Basters, A., Hess, D., Iesmantavicius, V., Lampersberger, L., Flemr, M., Pandey, A., Thoma, N.H.*, et al.* (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature *557*, 739-743.

Ozsolak, F., Kapranov, P., Foissac, S., Kim, S.W., Fishilevich, E., Monaghan, A.P., John, B., and Milos, P.M. (2010). Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. Cell *143*, 1018-1029.

Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol Cell Biol *7*, 4557-4559.

Pare, J.M., and Sullivan, C.S. (2014). Distinct antiviral responses in pluripotent versus differentiated cells. PLoS Pathog *10*, e1003865.

Parker, R. (2012). RNA Degradation in Saccharomyces cerevisae. Genetics *191*, 671-702.

Pashler, A.L., Towler, B.P., Jones, C.I., and Newbury, S.F. (2016). The roles of the exoribonucleases DIS3L2 and XRN1 in human disease. Biochem Soc T *44*, 1377-1384.

Pelechano, V., and Alepuz, P. (2017). eIF5A facilitates translation termination globally and promotes the elongation of many non polyproline-specific tripeptide sequences. Nucleic Acids Res *45*, 7326-7338.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. Cell *161*, 1400-1412.

Penheiter, K.L., Washburn, T.M., Porter, S.E., Hoffman, M.G., and Jaehning, J.A. (2005). A posttranscriptional role for the yeast Paf1-RNA polymerase II complex is revealed by identification of primary targets. Mol Cell *20*, 213-223.

Pérez-Ortín, J.E., Alepuz, P., Chávez, S., and Choder, M. (2013). Eukaryotic mRNA Decay: Methodologies, Pathways, and Links to Other Stages of Gene Expression. J Mol Biol *425*, 3750-3775.

Pirouz, M., Du, P., Munafo, M., and Gregory, R.I. (2016). Dis3l2-Mediated Decay Is a Quality Control Pathway for Noncoding RNAs. Cell Rep *16*, 1861-1873.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U., Pestova, T.V., and Pisarev, A.V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. Embo j *30*, 1804-1817.

Pokholok, D.K., Hannett, N.M., and Young, R.A. (2002). Exchange of RNA polymerase II initiation and elongation factors during gene expression in vivo. Mol Cell *9*, 799-809.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R.*, et al.* (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111-1124.

Radhakrishnan, A., Chen, Y.H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132.e129.

Rajavel, K.S., and Neufeld, E.F. (2001). Nonsense-mediated decay of human HEXA mRNA. Mol Cell Biol *21*, 5512-5519.

Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol Cell Biol *27*, 3612-3624.

Rissland, O.S., and Norbury, C.J. (2009). Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol *16*, 357–359.

Robb, G.B., Brown, K.M., Khurana, J., and Rana, T.M. (2005). Specific and potent RNAi in the nucleus of human cells. Nat Struct Mol Biol *12*, 133-137.

Robinson, S.R., Oliver, A.W., Chevassut, T.J., and Newbury, S.F. (2015). The 3' to 5' Exoribonuclease DIS3: From Structure and Mechanisms to Biological Functions and Role in Human Disease. Biomolecules *5*, 1515-1539.

Rozenblatt-Rosen, O., Nagaike, T., Francis, J.M., Kaneko, S., Glatt, K.A., Hughes, C.M., LaFramboise, T., Manley, J.L., and Meyerson, M. (2009). The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. Proc Natl Acad Sci USA *106*, 755-760.

Rufener, S.C., and Muhlemann, O. (2013). eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. Nat Struct Mol Biol *20*, 710-717.

Sadeghi, L., Prasad, P., Ekwall, K., Cohen, A., and Svensson, J.P. (2015). The Paf1 complex factors Leo1 and Paf1 promote local histone turnover to modulate chromatin states in fission yeast. EMBO Rep *16*, 1673-1687.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F.P., Dastidar, E.G., Sanchez-Rotunno, M., Arraiano, C.M., and van Hoof, A. (2009). The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat Struct Mol Biol *16*, 56-62.

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillere, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M.*, et al.* (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science *354*, 1431-1433.

Schmidt, M.J., West, S., and Norbury, C.J. (2011). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA *17*, 39-44.

Schmidt, S.A., Foley, P.L., Jeong, D.H., Rymarquis, L.A., Doyle, F., Tenenbaum, S.A., Belasco, J.G., and Green, P.J. (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic Acids Res *43*, 309-323.

Schneider, C., Leung, E., Brown, J., and Tollervey, D. (2009). The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res *37*, 1127-1140.

Schuller, A.P., Wu, C.C., Dever, T.E., Buskirk, A.R., and Green, R. (2017). eIF5A Functions Globally in Translation Elongation and Termination. Mol Cell *66*, 194-205.e195.

She, M., Decker, C.J., Svergun, D.I., Round, A., Chen, N., Muhlrad, D., Parker, R., and Song, H. (2008). Structural basis of dcp2 recognition and activation by dcp1. Mol Cell *29*, 337-349.

Shen, B., and Goodman, H.M. (2004). Uridine addition after microRNA-directed cleavage. Science *306*, 997.

Shi, X., Chang M Fau - Wolf, A.J., Wolf Aj Fau - Chang, C.H., Chang Ch Fau - Frazer-Abel, A.A., Frazer-Abel Aa Fau - Wade, P.A., Wade Pa Fau - Burton, Z.F., Burton Zf Fau - Jaehning, J.A., and Jaehning, J.A. (1997). Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme.

Shimada, Y., Mohn, F., and Buhler, M. (2016). The RNA-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts. Genes Dev *30*, 2571-2580.

Shoemaker, C.J., Eyler, D.E., and Green, R. (2010). Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. Science *330*, 369-372.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proc Natl Acad Sci USA *108*, E1392-1398.

Shoemaker, C.J., and Green, R. (2012). Translation drives mRNA quality control. Nature Structural &Amp; Molecular Biology *19*, 594.

Silva, A.L., Ribeiro, P., Inacio, A., Liebhaber, S.A., and Romao, L. (2008). Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. RNA *14*, 563-576.

Simmer, F., Buscaino, A., Kos-Braun, I.C., Kagansky, A., Boukaba, A., Urano, T., Kerr, A.R., and Allshire, R.C. (2010). Hairpin RNA induces secondary small interfering RNA synthesis and silencing in trans in fission yeast. EMBO Rep *11*, 112-118.

Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., and Zaher, H.S. (2014). An active role for the ribosome in determining the fate of oxidized mRNA. Cell Rep *9*, 1256-1264.

Simms, C.L., Yan, L.L., and Zaher, H.S. (2017). Ribosome Collision Is Critical for Quality Control during No-Go Decay. Mol Cell *68*, 361-373 e365.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS biology *6*, e111.

Singh, R.K., Gonzalez, M., Kabbaj, M.H., and Gunjan, A. (2012). Novel E3 ubiquitin ligases that regulate histone protein levels in the budding yeast Saccharomyces cerevisiae. PloS one *7*, e36295.

Sinkkonen, L., Hugenschmidt, T., Filipowicz, W., and Svoboda, P. (2010). Dicer is associated with ribosomal DNA chromatin in mammalian cells. PloS one *5*, e12175.

Sitron, C.S., Park, J.H., and Brandman, O. (2017). Asc1, Hel2, and Slh1 couple translation arrest to nascent chain degradation. RNA *23*, 798-810.

Slevin, Michael K., Meaux, S., Welch, Joshua D., Bigler, R., Miliani de Marval, Paula L., Su, W., Rhoads, Robert E., Prins, Jan F., and Marzluff, William F. (2014). Deep Sequencing Shows Multiple Oligouridylations Are Required for 3′ to 5′ Degradation of Histone mRNAs on Polyribosomes. Mol Cell *53*, 1020-1030.

Sohrabi-Jahromi, S., Hofmann, K.B., Boltendahl, A., Roth, C., Gressel, S., Baejen, C., Soeding, J., and Cramer, P. (2019). Transcriptome maps of general eukaryotic RNA degradation factors. eLife *8*, e47040.

Somers, J., Poyry, T., and Willis, A.E. (2013). A perspective on mammalian upstream open reading frame function. The international journal of biochemistry & cell biology *45*, 1690-1700.

Song, M.G., Bail, S., and Kiledjian, M. (2013). Multiple Nudix family proteins possess mRNA decapping activity. RNA *19*, 390-399.

Song, M.G., and Kiledjian, M. (2007). 3' Terminal oligo U-tract-mediated stimulation of decapping. RNA *13*, 2356-2365.

Song, M.G., Li, Y., and Kiledjian, M. (2010). Multiple mRNA decapping enzymes in mammalian cells. Mol Cell *40*, 423-432.

Squazzo, S.L., Costa, P.J., Lindstrom, D.L., Kumer, K.E., Simic, R., Jennings, J.L., Link, A.J., Arndt, K.M., and Hartzog, G.A. (2002). The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. EMBO J *21*, 1764-1774.
Staals, R.H., Bronkhorst, A.W., Schilders, G., Slomovic, S., Schuster, G., Heck, A.J., Raijmakers, R., and Pruijn, G.J. (2010). Dis3-like 1: a novel exoribonuclease associated with the human exosome. EMBO J *29*, 2358-2367.

Stalder, L., Heusermann, W., Sokol, L., Trojer, D., Wirz, J., Hean, J., Fritzsche, A., Aeschimann, F., Pfanzagl, V., Basselet, P.*, et al.* (2013). The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing. EMBO J *32*, 1115-1127.

Su, W., Slepenkov, S.V., Slevin, M.K., Lyons, S.M., Ziemniak, M., Kowalska, J., Darzynkiewicz, E., Jemielity, J., Marzluff, W.F., and Rhoads, R.E. (2013). mRNAs containing the histone 3' stem-loop are degraded primarily by decapping mediated by oligouridylation of the 3' end. RNA *19*, 1-16.

Sun, M., Schwalb, B., Pirkl, N., Maier, K.C., Schenk, A., Failmezger, H., Tresch, A., and Cramer, P. (2013). Global analysis of eukaryotic mRNA degradation reveals Xrn1-dependent buffering of transcript levels. Mol Cell *52*, 52-62.

Sundaramoorthy, E., Leonard, M., Mak, R., Liao, J., Fulzele, A., and Bennett, E.J. (2017). ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. Mol Cell *65*, 751-760.e754.

Suzuki, K., Shijuuku, T., Fukamachi, T., Zaunders, J., Guillemin, G., Cooper, D., and Kelleher, A. (2005). Prolonged transcriptional silencing and CpG methylation induced by siRNAs targeted to the HIV-1 promoter region. J RNAi Gene Silencing *1*, 66-78.

Svoboda, P. (2014). Renaissance of mammalian endogenous RNAi. FEBS Lett *588*, 2550-2556.
Synowsky, S.A., and Heck, A.J. (2008). The yeast Ski complex is a hetero-tetramer. Protein Sci *17*, 119-125.

Tan, D., Marzluff, W.F., Dominski, Z., and Tong, L. (2013). Structure of histone mRNA stem-loop, human stem-loop binding protein, and 3'hExo ternary complex. Science *339*, 318-321.

Tesina, P., Heckel, E., Cheng, J., Fromont-Racine, M., Buschauer, R., Kater, L., Beatrix, B., Berninghausen, O., Jacquier, A., Becker, T.*, et al.* (2019). Structure of the 80S ribosome-Xrn1 nuclease complex. Nat Struct Mol Biol *26*, 275-280.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol Cell *50*, 613-623.

Thandapani, P., Song, J., Gandin, V., Cai, Y., Rouleau, S.G., Garant, J.-M., Boisvert, F.-M., Yu, Z., Perreault, J.-P., Topisirovic, I.*, et al.* (2015). Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. eLife *4*, e06234.

Tharun, S., and Parker, R. (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol Cell *8*, 1075-1083.

Tippana, R., Chen, M.C., Demeshkina, N.A., Ferre-D'Amare, A.R., and Myong, S. (2019). RNA G-quadruplex is resolved by repetitive and ATP-dependent mechanism of DHX36. Nature communications *10*, 1855.

Tomecki, R., Kristiansen, M.S., Lykke-Andersen, S., Chlebowski, A., Larsen, K.M., Szczesny, R.J., Drazkowska, K., Pastula, A., Andersen, J.S., Stepien, P.P.*, et al.* (2010). The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. Embo j *29*, 2342-2357.

Tomson, B.N., and Arndt, K.M. (2013). The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochim Biophys Acta *1829*, 116-126.

Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., and Inada, T. (2012). Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. Mol Cell *46*, 518-529.

Tuck, A.C., Rankova, A., Arpat, A.B., Liechti, L.A., Hess, D., Iesmantavicius, V., Castelo-Szekely, V., Gatfield, D., and Buhler, M. (2020). Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation. Mol Cell *77*, 1222-1236.e1213.

Tuck, Alex C., and Tollervey, D. (2013). A Transcriptome-wide Atlas of RNP Composition Reveals Diverse Classes of mRNAs and lncRNAs. Cell *154*, 996-1009.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. EMBO J *21*, 1427-1436.

Tucker, M., Valencia-Sanchez, M.A., Staples, R.R., Chen, J., Denis, C.L., and Parker, R. (2001). The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. Cell *104*, 377-386.

Uchida, N., Hoshino, S., and Katada, T. (2004). Identification of a human cytoplasmic poly(A) nuclease complex stimulated by poly(A)-binding protein. J Biol Chem *279*, 1383-1391.

Unterholzner, L., and Izaurralde, E. (2004). SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Mol Cell *16*, 587-596.

Ustianenko, D., Pasulka, J., Feketova, Z., Bednarik, L., Zigackova, D., Fortova, A., Zavolan, M., and Vanacova, S. (2016). TUT-DIS3L2 is a mammalian surveillance pathway for aberrant structured non-coding RNAs. Embo j *35*, 2179-2191.

van den Elzen, A.M., Schuller, A., Green, R., and Seraphin, B. (2014). Dom34-Hbs1 mediated dissociation of inactive 80S ribosomes promotes restart of translation after stress. Embo j *33*, 265-276.

van Hoof, A., Frischmeyer, P.A., Dietz, H.C., and Parker, R. (2002). Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. Science *295*, 2262-2264.

van Hoof, A., Staples, R.R., Baker, R.E., and Parker, R. (2000). Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. Mol Cell Biol *20*, 8230-8243.

Van Oss, S.B., Cucinotta, C.E., and Arndt, K.M. (2017). Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. Trends Biochem Sci *42*, 788-798.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Wade, P.A., Werel, W., Fentzke, R.C., Thompson, N.E., Leykam, J.F., Burgess, R.R., Jaehning, J.A., and Burton, Z.F. (1996). A novel collection of accessory factors associated with yeast RNA polymerase II. Protein Expr Purif *8*, 85-90.

Wahle, E., and Winkler, G.S. (2013). RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. Biochim Biophys Acta *1829*, 561-570.

Wang, L., Lewis, M.S., and Johnson, A.W. (2005). Domain interactions within the Ski2/3/8 complex and between the Ski complex and Ski7p. RNA *11*, 1291-1302.

Wang, R., Wang, J., Paul, A.M., Acharya, D., Bai, F., Huang, F., and Guo, Y.L. (2013). Mouse embryonic stem cells are deficient in type I interferon expression in response to viral infections and double-stranded RNA. J Biol Chem *288*, 15926-15936.

Wang, Z., Jiao, X., Carr-Schmid, A., and Kiledjian, M. (2002). The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc Natl Acad Sci USA *99*, 12663-12668.

Webster, M.W., Chen, Y.-H., Stowell, J.A.W., Alhusaini, N., Sweet, T., Graveley, B.R., Coller, J., and Passmore, L.A. (2018). mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. Molecular cell *70*, 1089-1100.e1088.

Weick, E.M., Puno, M.R., Januszyk, K., Zinder, J.C., DiMattia, M.A., and Lima, C.D. (2018). Helicase-Dependent RNA Decay Illuminated by a Cryo-EM Structure of a Human Nuclear RNA Exosome-MTR4 Complex. Cell *173*, 1663-1677.e1621.

Weng, Y., Czaplinski, K., and Peltz, S.W. (1996). Identification and characterization of mutations in the UPF1 gene that affect nonsense suppression and the formation of the Upf protein complex but not mRNA turnover. Mol Cell Biol *16*, 5491-5506.

Weren, R.D., Venkatachalam, R., Cazier, J.B., Farin, H.F., Kets, C.M., de Voer, R.M., Vreede, L., Verwiel, E.T., van Asseldonk, M., Kamping, E.J.*, et al.* (2015). Germline deletions in the tumour suppressor gene FOCAD are associated with polyposis and colorectal cancer development. J Pathol *236*, 155-164.

White, E., Schlackow, M., Kamieniarz-Gdula, K., Proudfoot, N.J., and Gullerova, M. (2014). Human nuclear Dicer restricts the deleterious accumulation of endogenous double-stranded RNA. Nat Struct Mol Biol *21*, 552-559.

Whitfield, M.L., Zheng, L.X., Baldwin, A., Ohta, T., Hurt, M.M., and Marzluff, W.F. (2000). Stem-loop binding protein, the protein that binds the 3' end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. Mol Cell Biol *20*, 4188-4198.

Williams, A.S., and Marzluff, W.F. (1995). The sequence of the stem and flanking sequences at the 3' end of histone mRNA are critical determinants for the binding of the stem-loop binding protein. Nucleic Acids Res *23*, 654-662.

Winz, M.L., Peil, L., Turowski, T.W., Rappsilber, J., and Tollervey, D. (2019). Molecular interactions between Hel2 and RNA supporting ribosome-associated quality control. Nature communications *10*, 563.

Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. J Biol Chem *278*, 34739-34742.

Yamashita, A., Chang, T.C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.Y., and Shyu, A.B. (2005). Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. Nat Struct Mol Biol *12*, 1054-1063.

Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., and Ohno, S. (2001). Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. Genes Dev *15*, 2215-2228.

Yang, X.C., Purdy, M., Marzluff, W.F., and Dominski, Z. (2006). Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA. J Biol Chem *281*, 30447-30454.

Yang, Y., Li, W., Hoque, M., Hou, L., Shen, S., Tian, B., and Dynlacht, B.D. (2016). PAF Complex Plays Novel Subunit-Specific Roles in Alternative Cleavage and Polyadenylation. PLoS Genet *12*, e1005794.

Yangyuoru, P.M., Bradburn, D.A., Liu, Z., Xiao, T.S., and Russell, R. (2018). The G-quadruplex (G4) resolvase DHX36 efficiently and specifically disrupts DNA G4s via a translocation-based helicase mechanism. J Biol Chem *293*, 1924-1932.

Yi, H., Park, J., Ha, M., Lim, J., Chang, H., and Kim, V.N. (2018). PABP Cooperates with the CCR4-NOT Complex to Promote mRNA Deadenylation and Block Precocious Decay. Mol Cell *70*, 1081-1088 e1085.

Yu, M., Yang, W., Ni, T., Tang, Z., Nakadai, T., Zhu, J., and Roeder, R.G. (2015). RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. Science *350*, 1383-1386.

Zabolotskaya, M.V., Grima, D.P., Lin, M.D., Chou, T.B., and Newbury, S.F. (2008). The 5'-3' exoribonuclease Pacman is required for normal male fertility and is dynamically localized in cytoplasmic particles in Drosophila testis cells. The Biochemical journal *416*, 327-335.

Zeng, Y., and Cullen, B.R. (2002). RNA interference in human cells is restricted to the cytoplasm. RNA *8*, 855-860.

Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019). A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. Embo j *38*, e100640.

Zhou, W., Lu, Q., Li, Q., Wang, L., Ding, S., Zhang, A., Wen, X., Zhang, L., and Lu, C. (2017). PPR-SMR protein SOT1 has RNA endonuclease activity. Proc Natl Acad Sci USA *114*, E1554-e1563.

Zhu, B., Mandal, S.S., Pham, A.D., Zheng, Y., Erdjument-Bromage, H., Batra, S.K., Tempst, P., and Reinberg, D. (2005a). The human PAF complex coordinates transcription with events downstream of RNA synthesis. Genes Dev *19*, 1668-1673.

Zhu, B., Zheng, Y., Pham, A.D., Mandal, S.S., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2005b). Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Mol Cell *20*, 601-611.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev *31*, 88-100.

Zong, J., Yao, X., Yin, J., Zhang, D., and Ma, H. (2009). Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene *447*, 29-39.

Zund, D., Gruber, A.R., Zavolan, M., and Muhlemann, O. (2013). Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3' UTRs. Nat Struct Mol Biol *20*, 936-943.

# Appendix

**Mammalian RNA decay pathways are highly specialized and widely linked to translation**

Alex Charles Tuck[1,4], Aneliya Rankova[1,4], Alaaddin Bulak Arpat[3], Luz Angelica Liechti[3], Daniel Hess[1], Vytautas Iesmantavicius[1], Violeta Castelo-Szekely[3], David Gatfield[3] and Marc Bühler[1,2,5]

*Molecular cell, 2020*

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

[2]University of Basel, Petersplatz 10, 4003 Basel, Switzerland

[3]Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland

[4]These authors contributed equally

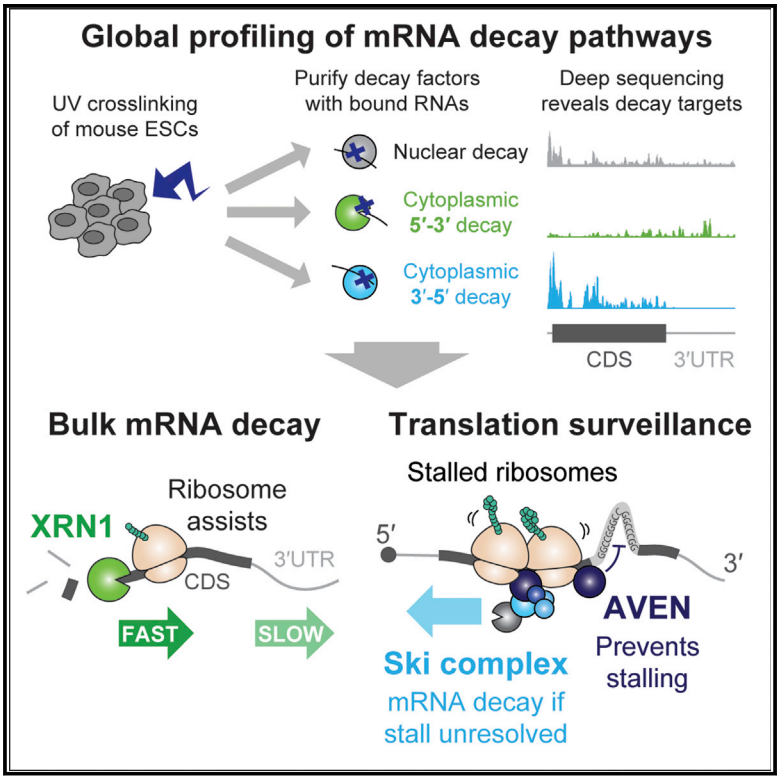
[5]Lead contact

Article

Molecular Cell

# Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation

## Graphical Abstract




## Authors

Alex Charles Tuck, Aneliya Rankova, Alaaddin Bulak Arpat, ..., Violeta Castelo-Szekely, David Gatfield, Marc Bühler

## Correspondence

marc.buehler@fmi.ch

## In Brief

Tuck, Rankova, et al. globally profile mRNA decay and aberrant translation events in mouse embryonic stem cells, finding that mRNA decay pathways perform specialized roles rather than acting redundantly. They uncover widespread crosstalk between mRNA decay and translation and identify AVEN as a factor that counteracts ribosome stalling.

## Highlights

- Global profiling of mRNA decay pathways and aberrant translation events in mESCs
- XRN1 mediates mRNA turnover, whereas SKIV2L acts widely in translation surveillance
- AVEN interacts with ribosomes and the Ski complex and counteracts ribosome stalling
- Histone mRNAs, uORFs, and small ORFs are key targets of SKIV2L and AVEN

Tuck et al., 2020, Molecular Cell 77, 1222–1236
March 19, 2020 
https://doi.org/10.1016/j.molcel.2020.01.007

CellPress

CellPress

Molecular Cell
Article

# Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation

Alex Charles Tuck,[1,4] Aneliya Rankova,[1,4] Alaaddin Bulak Arpat,[3] Luz Angelica Liechti,[3] Daniel Hess,[1] Vytautas Iesmantavicius,[1] Violeta Castelo-Szekely,[3] David Gatfield,[3] and Marc Bühler[1,2,5,*]

[1]Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland
[2]University of Basel, Petersplatz 10, 4003 Basel, Switzerland
[3]Center for Integrative Genomics, University of Lausanne, 1015 Lausanne, Switzerland
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: marc.buehler@fmi.ch
https://doi.org/10.1016/j.molcel.2020.01.007

## SUMMARY

**RNA decay is crucial for mRNA turnover and surveillance and misregulated in many diseases. This complex system is challenging to study, particularly in mammals, where it remains unclear whether decay pathways perform specialized versus redundant roles. Cytoplasmic pathways and links to translation are particularly enigmatic. By directly profiling decay factor targets and normal versus aberrant translation in mouse embryonic stem cells (mESCs), we uncovered extensive decay pathway specialization and crosstalk with translation. XRN1 (5′-3′) mediates cytoplasmic bulk mRNA turnover whereas SKIV2L (3′-5′) is universally recruited by ribosomes, tackling aberrant translation and sometimes modulating mRNA abundance. Further exploring translation surveillance revealed AVEN and FOCAD as SKIV2L interactors. AVEN prevents ribosome stalls at structured regions, which otherwise require SKIV2L for clearance. This pathway is crucial for histone translation, upstream open reading frame (uORF) regulation, and counteracting ribosome arrest on small ORFs. In summary, we uncovered key targets, components, and functions of mammalian RNA decay pathways and extensive coupling to translation.**

## INTRODUCTION

RNA decay ensures steady-state mRNA expression, eliminates aberrant transcripts, and remodels the transcriptome upon changing conditions (Bresson et al., 2017; Pérez-Ortín et al., 2013; Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013). In the nucleus, mRNAs are mainly degraded 3′–5′ by the exosome complex, assisted by factors including the helicase Mtr4 (MTR4) (Kilchert et al., 2016; LaCava et al., 2005; Mitchell et al., 1997; Schmid and Jensen, 2018). Cytoplasmic mRNA turnover is initiated by poly(A) tail removal and proceeds via 3′–5′ exoribonucleolysis by the exosome or decapping followed by 5′–3′ degradation by the exoribonuclease Xrn1 (XRN1) (Hsu and Stevens, 1993; Łabno et al., 2016; Parker, 2012; Zinder and Lima, 2017). Cytoplasmic exosome activity requires the Ski complex (Anderson and Parker, 1998), comprising the scaffold Ski3 (TTC37), two copies of Ski8 (WDR61), and the helicase Ski2 (SKIV2L). Ski2, like its homolog Mtr4, unwinds RNA and channels it to the exosome (Halbach et al., 2013). Many pathologies are linked to dysregulation of these factors. For example, XRN1 is downregulated in osteosarcoma (Pashler et al., 2016), exosome mutations are linked to cancer (Robinson et al., 2015) and neurological disorders (Morton et al., 2018), and Ski complex impairment causes a congenital bowl disorder (Fabre et al., 2012; Hartley et al., 2010).

The complexity of RNA decay makes it hard to study and fundamental questions remain. For example, do decay pathways act redundantly or target specific transcripts? If the latter, how is specificity achieved, and what advantage does it confer? Analyses of *S. cerevisiae* mutants suggest that Xrn1 contributes more than the exosome to cytoplasmic turnover (Parker, 2012). However, compensation between decay pathways and secondary effects make it unclear whether this reflects the physiological situation. Furthermore, higher eukaryotes have extra factors and pathways, including 3′ uridyltransferases acting in cytoplasmic decay (Łabno et al., 2016; Lim et al., 2014) and diverse MTR4-containing nuclear exosome adaptor complexes (Lubas et al., 2011; Meola et al., 2016).

A further challenge is that RNA decay is coupled to other RNA life cycle events. For example, the nuclear exosome is recruited during transcription to remove early termination products, introns, or full-length mRNAs (Kilchert et al., 2016). In the cytoplasm, there is crosstalk between translation and RNA decay, epitomized by surveillance pathways targeting mRNAs with premature termination codons (nonsense-mediated decay [NMD]), translational roadblocks (no-go decay [NGD]), or no stop codon (nonstop decay [NSD]) (Roy and Jacobson, 2013). A key event is mRNA cleavage at stalled ribosomes, which generates 5′ and 3′ RNA fragments that are cleared by the exosome and Xrn1 (Gatfield and Izaurralde, 2004; Ghosh and Jacobson, 2010; Guydosh and Green, 2017). Coupling between translation and degradation could be widespread and extend beyond surveillance (Ibrahim et al., 2018). In support of this, Xrn1 can act co-translationally (Hu et al., 2009; Pelechano et al., 2015), and

structures capture the yeast Ski complex or Xrn1 bound to ribosomes (Schmidt et al., 2016; Tesina et al., 2019). There is intense interest in understanding whether decay factor interactions with the ribosome are conserved in higher eukaryotes, the functional relevance, and whether this constitutes a major decay route.

Here, we address key questions about mammalian mRNA decay. First, what are the physiological targets of major decay pathways? Second, focusing on cytoplasmic decay, to what extent is this coupled to translation, and what factors influence this? To reveal direct, physiological targets of decay factors, we used crosslinking and analysis of cDNAs (CRAC) to compare the transcriptome-wide interactions of XRN1, SKIV2L, and MTR4 in mouse embryonic stem cells (mESCs). Our data suggest that most mRNA turnover occurs via the 5′–3′ pathway, but some mRNAs (particularly those encoding histones) depend on cytoplasmic 3′–5′ decay. We find that SKIV2L and XRN1 directly bind ribosomes, and translation appears to assist bulk mRNA turnover by XRN1. Strikingly, SKIV2L is specifically and pervasively recruited to ribosome-occupied regions, suggesting it acts exclusively in translation-associated mRNA surveillance. Our data reveal triggers of ribosome stalling and SKIV2L recruitment, which we explore by globally mapping stalled ribosomes. Proteomic analyses identify the RNA-binding factor AVEN and uncharacterized protein FOCAD as Ski complex interactors. We observe AVEN binding to GC-rich RNAs predicted to be structured and increased SKIV2L binding, decay, and ribosome stalling at these regions upon *Aven* knockout. We conclude that AVEN and SKIV2L cooperate to counteract aberrant translation, with AVEN preventing ribosome stalls at structured regions and SKIV2L eliminating transcripts if these events accumulate. The AVEN-SKIV2L pathway acts on diverse substrates, including histone mRNAs, upstream open reading frames (uORFs), and small ORF (sORF)-containing RNAs. In summary, we uncover specialization between mammalian RNA decay pathways and widespread crosstalk with translation and establish SKIV2L and AVEN as components of a universal translation surveillance program.

## RESULTS

### Mammalian RNA Decay Pathways Target Distinct Transcripts

To examine the specificity of RNA decay pathways (Figure 1A), we applied the CRAC approach to SKIV2L, XRN1, and MTR4 in mESCs (Granneman et al., 2009; Tuck et al., 2018). After endogenously 3xFLAG-Avi tagging these proteins (Figure S1A; Table S1) (Flemr and Bühler, 2015), we crosslinked cells with UV (254 nm) to fix protein-RNA interactions, purified ribonucleoproteins (RNPs) under denaturing conditions, performed a limited RNase digestion, and sequenced the RNA fragments (Figure 1B). We performed five or six technical replicates (including three published MTR4 datasets; Table S2). Global comparison of mRNAs bound by SKIV2L, MTR4, and XRN1 using principal-component analysis (PCA) or correlation coefficients revealed highly reproducible differences (Figures 1C, 1D, and S1B; Table S3). To explore the specificity of individual transcripts, we used t-distributed stochastic neighbor embedding (t-SNE) to arrange mRNAs by relative binding to the three proteins (Figure 1E). Although some transcripts bound similarly to SKIV2L, XRN1, and MTR4 (e.g., *Trim28*; Figure 1F), others had a clear preference (e.g., *Sfpq* or *Pim3*; Figure 1F), suggesting that for many transcripts, one decay route dominates. Furthermore, functionally related mRNAs shared binding preferences (Figures 1E and S1D) (e.g., histone mRNAs bound abundantly to SKIV2L).

As XRN1-dependent 5′–3′ decay is assumed to be the main determinant of RNA half-life and steady-state abundance we were intrigued by transcripts bound highly by SKIV2L (e.g., Figure 1G). SKIV2L assists the exosome in 3′–5′ decay, and a co-immunoprecipitation confirmed that 3xFLAG-Avi-tagged SKIV2L interacts with the cytoplasmic exosome component DIS3L (Figure S1C). We therefore suspected that highly SKIV2L-bound mRNAs are degraded in a 3′–5′ SKIV2L-dependent manner. To test this, we generated *Skiv2l*$^{fl/fl}$ conditional knockout cells by integrating *loxP* sites into introns 10 and 17 in CreERT2-expressing mESCs (Flemr and Bühler, 2015) (Figure 1H). We treated these cells with 4-hydroxytamoxifen (4OHT) to induce *loxP* recombination and production of truncated SKIV2L without a catalytic domain (Figures 1H and S1E) and profiled gene expression by RNA sequencing (RNA-seq) (Table S3). There were many changes after 6 days of 4OHT treatment, but these did not correlate with SKIV2L CRAC (Figure 1I, right) so are likely indirect effects. Conversely, after 4 days of 4OHT treatment, transcript upregulation correlated with SKIV2L CRAC (Figure 1I, left, and Figure S1F). Measuring transcriptome-wide half-lives following transcription shut off by actinomycin D confirmed that highly SKIV2L-bound transcripts are stabilized upon *Skiv2l* knockout, exemplified by replication-dependent histone mRNAs (Figure S1G; Table S3). Some stabilized SKIV2L targets (e.g., *Calr* and *Pdia4*; Figure S1G) did not increase in abundance (Figure 1I), suggesting that cells partially compensate for the loss of SKIV2L. Of note, high-confidence SKIV2L targets (Figure S1G) were expressed at wild-type (WT) levels in our tagged cell lines, confirming that tagged SKIV2L is functional (Figure S1H). We conclude that SKIV2L-dependent 3′–5′ decay contributes to the steady-state abundance of a subset of mRNAs, including most replication-dependent histone mRNAs. Our approach thus reveals physiological targets of mRNA decay pathways.

### Cytoplasmic RNA Decay Is Widely Influenced by Translation

As cytoplasmic decay pathways are less well studied, we now focused on XRN1 and SKIV2L. A key question is to what extent they interact with translation. Remarkably, CRAC reads mapping to ribosomal RNA revealed specific, reproducible binding of SKIV2L and XRN1 to the 40S subunit mRNA entry and exit regions (Figure 2A), resembling yeast structures (Schmidt et al., 2016; Tesina et al., 2019). Therefore, SKIV2L and XRN1 ribosome interactions are conserved to mammals and occur in unperturbed cells.

To explore whether SKIV2L and/or XRN1 activity is widely coupled to translation, we examined binding across individual mRNAs (e.g., Figures 1F and 1G). SKIV2L binding was strongly biased toward regions occupied by ribosomes, i.e., 5′ UTRs, coding sequences (CDSs), and uORFs (e.g., *Ifrd1*; Figure 1G).

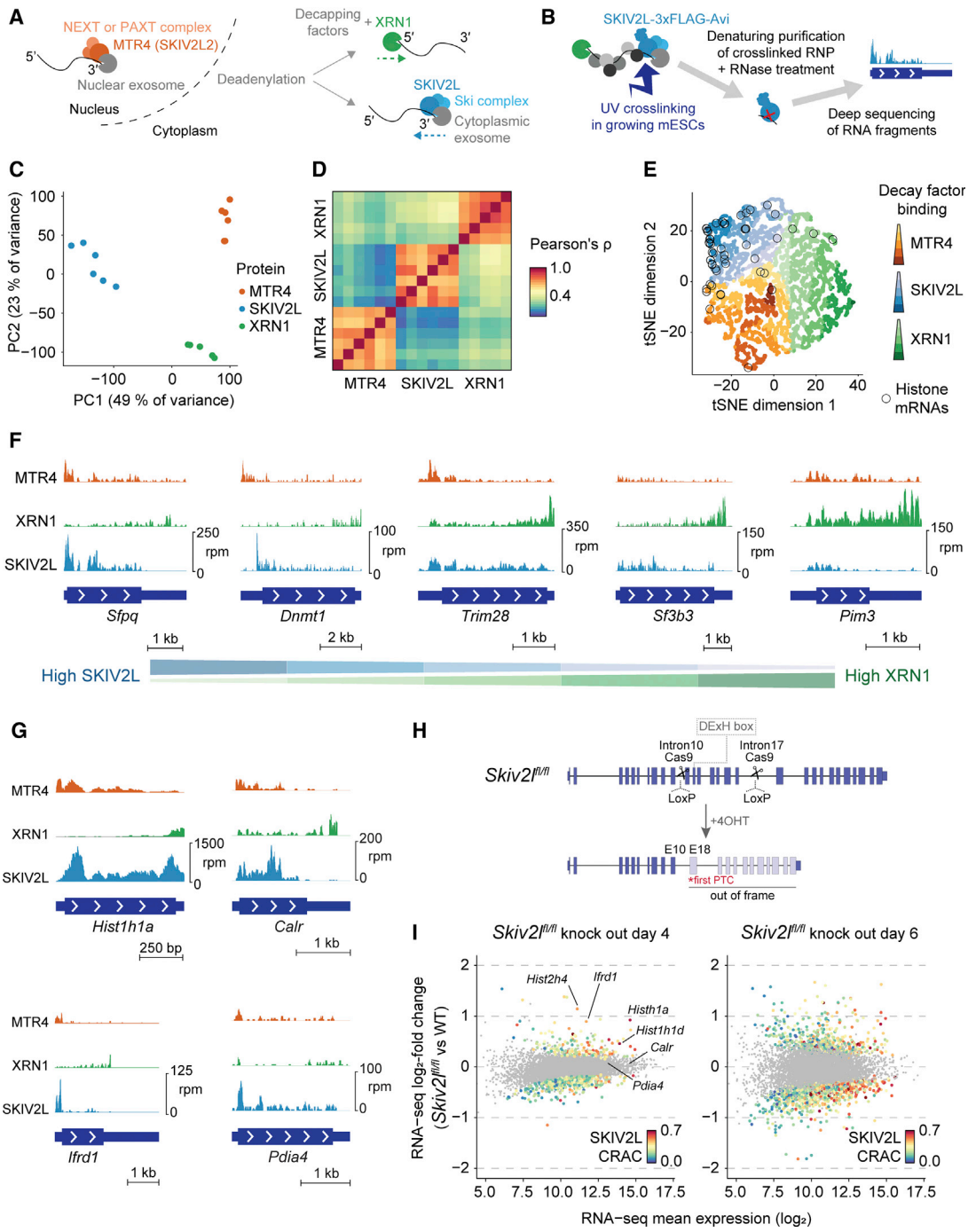



**Figure 1. Mammalian mRNA Decay Pathways Target Distinct Transcripts**

(A) RNA decay pathways.

(B) CRAC outline.

(C and D) PCA (C) and correlation matrix (D) based on decay factor binding (CRAC counts) to mRNAs. Replicates correspond to separate experiments for the same cell line.

(E) t-SNE representation of mRNAs based on relative binding to MTR4, SKIV2L, and XRN1.

(F and G) CRAC coverage across individual mRNAs. Transcripts in (F) illustrate different XRN1:SKIV2L ratios, whereas (G) depicts transcripts highlighed in panel (I).

(H) Conditional knockout strategy for *Skiv2l*.

(I) Differential expression analysis for *Skiv2l* knockout for the mRNAs in (E), with significantly changing transcripts (DESeq2 padj < 0.05) colored by SKIV2L binding (as in E).

See also Figure S1 and Tables S1, S2, and S3.

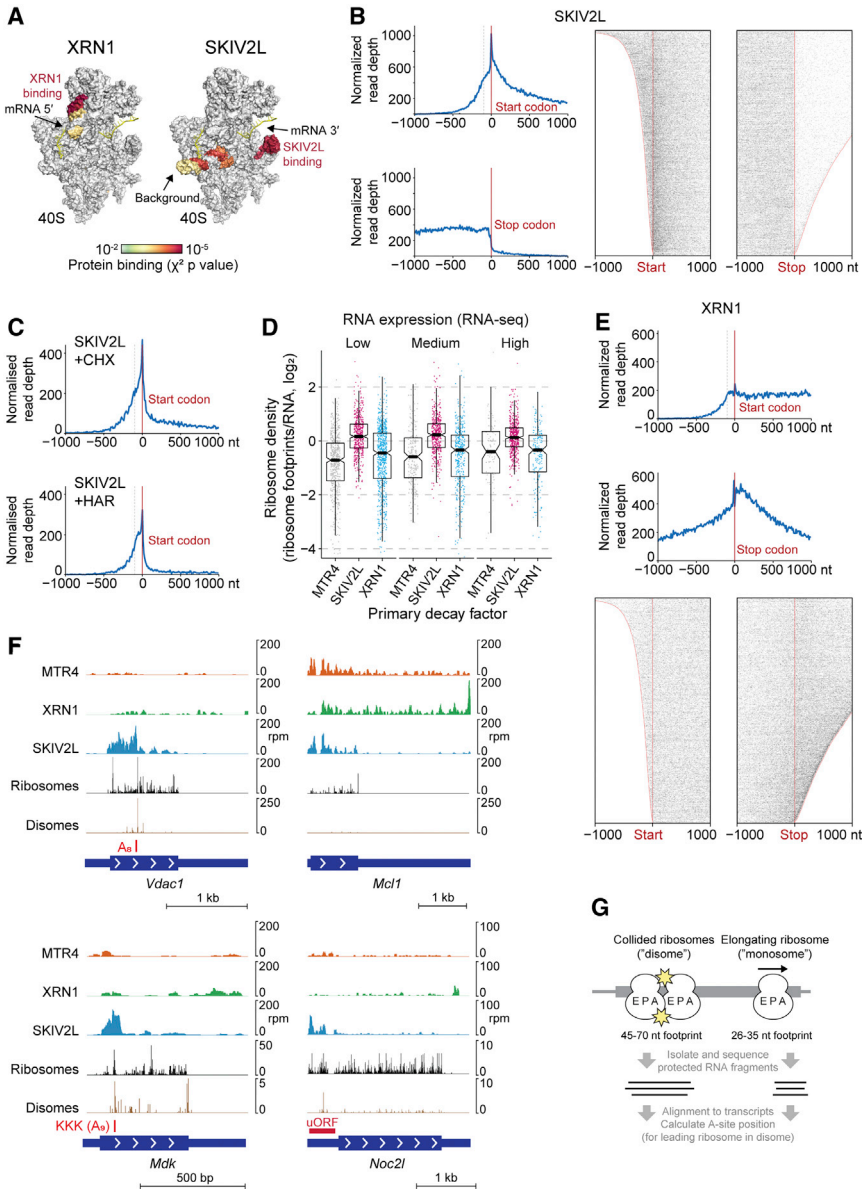



**Figure 2. Cytoplasmic mRNA Decay Is Widely Influenced by Translation**

(A) CRAC signal for SKIV2L and XRN1 on the ribosomal 40S subunit, based on the mouse rRNA sequence and human structure (Khatter et al., 2015). Significantly bound regions are colored by $\chi^2$ p value, and the mRNA path (yellow) is taken from Schmidt et al. (2016).

(B and C) CRAC signal for SKIV2L around start and stop codons, summed (left) or for individual mRNAs (right). Data in (B) correspond to untreated cells, whereas those in (C) correspond to 30-min cycloheximide or harringtonine treatment.

(D) Ribosome densities for mRNAs grouped by expression and most abundantly bound decay factor (defined in Figure 1E).

(E) XRN1 CRAC signal around start and stop codons.

(F) CRAC, monosome, and disome profiling for individual mRNAs.

(G) Monosome and disome profiling approach.

See also Figure S2 and Table S4.

Global analysis of binding around start and stop codons (Figure 2B) revealed this pattern is universal. Treating cells with translation inhibitors led to a redistribution of SKIV2L binding (Figure 2C) that parallels changes in ribosome occupancy, confirming that active translation directs SKIV2L binding. Harringtonine blocks translation post-initiation to deplete ribosomes from CDS regions, where we observed loss of SKIV2L binding. In contrast, cycloheximide stalls elongating ribosomes, leading to queuing and initiation upstream of the canonical start codon (Kearse et al., 2019). Consistently, SKIV2L accumulated in 5′ UTRs (Figure 2C). Further supporting the role of ribosomes in recruiting SKIV2L, we found that SKIV2L CRAC correlates with the number of ribosomes on a transcript, which we measured by ribosome profiling (Figure 2D; Table S4). We conclude that SKIV2L is specifically and universally recruited to translated regions via ribosome interactions.

In contrast to SKIV2L, XRN1 bound the full length of mRNAs, consistent with its major role being in bulk mRNA turnover. Strong XRN1 enrichment in 3′ UTRs (Figure 2E) supports a model where XRN1 follows the last translating ribosome, which helps remove obstacles. In the 3′ UTR, XRN1 may stall at RNA structures or protein-bound sites. The pattern of XRN1 binding around the stop codon is less well defined than that of SKIV2L, supporting this looser relationship with the ribosome. We conclude that both cytoplasmic decay pathways are widely influenced by translation, but only XRN1 degrades full-length mRNAs.

## SKIV2L Functions in Universal Translation Surveillance

We next sought to identify translation events leading to SKIV2L recruitment. Unlike the relatively even ribosome profiling coverage across mRNAs, SKIV2L CRAC signal was enriched at specific sites (e.g., Figure 2F). Cytoplasmic 3′–5′ decay acts in many surveillance pathways (e.g., NMD, NGD, and NSD), so we suspected that SKIV2L peaks reflect RNA features that arrest or stall ribosomes. Endonucleolytic cleavage at ribosome stall sites (D'Orazio et al., 2019; Gatfield and Izaurralde, 2004; Glover et al., 2019; Guydosh and Green, 2017) may enable SKIV2L to engage the 3′ end of the upstream fragment (Schmidt et al., 2016). Consistent with this, some SKIV2L-bound RNA fragments had non-templated 3′ U-tails (Figure S2A). Uridylation facilitates mRNA degradation by XRN1, DIS3L2, and the exosome (Lim et al., 2014; Slevin et al., 2014) and may act as a landing pad for SKIV2L. The U-tails confirm that SKIV2L binds cleaved RNAs. We also found U-tails on XRN1-bound RNA fragments, consistent with 5′–3′ and 3′–5′ pathways being able to act on a single

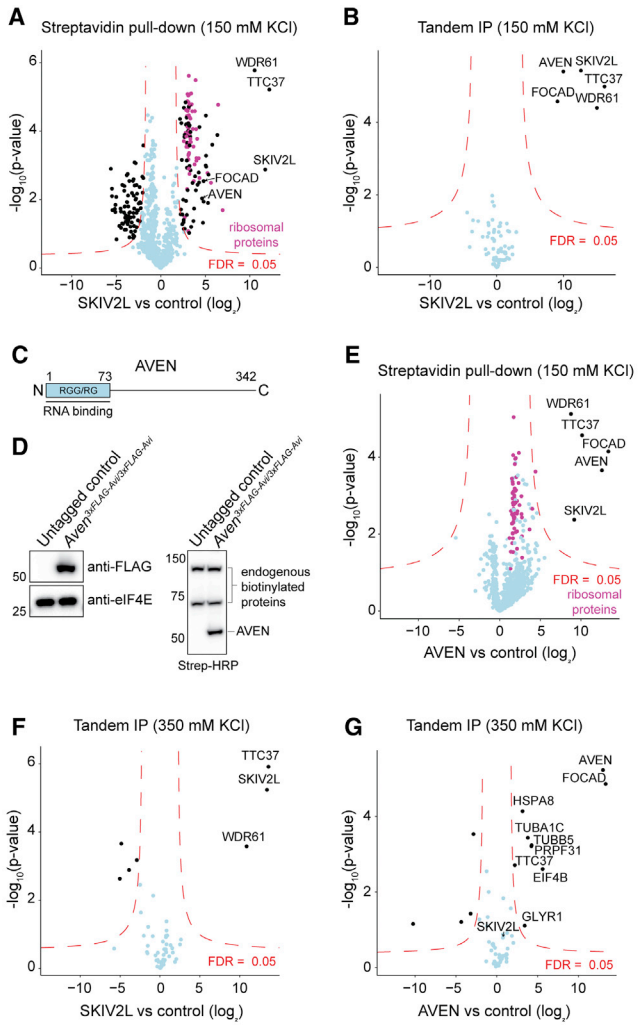



**Figure 3. AVEN and FOCAD are Ski Complex Interactors**

(A and B) Mass spectrometry (MS) of streptavidin (A) or tandem FLAG-streptavidin (B) purification of 3xFLAG-Avi-SKIV2L.

(C) Mouse AVEN protein.

(D) Western blot analysis of endogenously tagged *Aven*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ expression and biotinylation.

(E) MS of streptavidin purification of 3xFLAG-Avi-AVEN.

(F and G) MS of tandem FLAG-streptavidin purification of 3xFLAG-Avi-SKIV2L (F) and 3xFLAG-Avi-AVEN (G) using high salt. All experiments include RNase treatment, three technical replicates, and untagged mESCs as a control. FDR, false discovery rate.

See also Table S5.

mRNA and as reported by studies of yeast antiviral activity (Widner and Wickner, 1993) and for histone mRNAs (Mullen and Marzluff, 2008).

We reasoned that 3′ ends of SKIV2L-bound RNA fragments should reveal endogenous triggers of ribosome stalling. Indeed, 3′ ends were enriched at specific codon pairs, including those encoding lysine-lysine or proline (Figure S2B). Enrichment at proline codons was weak but had a clear frame preference, corroborating reports that proline in the nascent peptide triggers stalling (Ingolia et al., 2011; Pavlov et al., 2009). Examining longer codon runs, SKIV2L binding was elevated at poly-proline, -lysine, -glutamate, -aspartate, and -arginine (Figure S2C; *Mdk* in Figure 2F). These preferences resemble codons reported to stall ribosomes based upon mESC ribosome profiling (Ingolia et al., 2011). As SKIV2L peaks occurred at purine-rich codon runs, we suspected that for these, the RNA sequence is more important than the amino acid. Examining runs of $\geq$12 purines, SKIV2L enrichment was equivalent at lysine-rich and lysine-poor sequences but more pronounced at A-rich than G-rich sequences (Figure S2D). This suggests that A-rich sequences trigger ribosome stalling and SKIV2L surveillance, as exemplified by *Vdac1* and *Mdk* (Figure 2F, red boxes), and agrees with a reporter-based study (Arthur et al., 2015). XRN1 showed slight enrichment at some of these sites (Figures S2C and S2D), likely reflecting a minor role in surveillance.

To verify that SKIV2L-bound sites reflect ribosome stalls, we used a new method (disome profiling) to map collided ribosome pairs (disomes) (Figure 2G) (Arpat et al., 2019). Disomes form at ribosome stall sites (Ikeuchi et al., 2019; Juszkiewicz et al., 2018; Simms et al., 2017) and can be identified from 45- to 70-nt protected RNA fragments (Arpat et al., 2019; Guydosh and Green, 2014). We also performed standard ribosome profiling (monosome profiling). We calculated A-site positions of monosomes and leading ribosomes in disomes (Figure 2G). This revealed disome enrichments over codon and sequence runs (Figures S2C and S2D) and individual sites (e.g., *Vdac1*, *Mdk*, and *Noc2l*; Figure 2F) with elevated SKIV2L binding, confirming these reflect ribosome stalling. In some cases (e.g., polyproline; Figure S2C), the disome signal was stronger than the SKIV2L CRAC signal. This suggests that some stalls potently trigger RNA cleavage, but others (e.g., polyproline) are resolved without mRNA decay.

We conclude that although SKIV2L and XRN1 can target the same transcript, their roles are highly specialized. XRN1 mediates bulk mRNA decay, with a minor surveillance role, whereas SKIV2L responds exclusively to aberrant translation.

## AVEN and FOCAD Are Ski Complex Interactors

We next wondered if SKIV2L is recruited solely by ribosome and mRNA interactions or if other factors participate. MTR4 is targeted by adaptor proteins, so analogous Ski complex adapters could also exist. To identify SKIV2L interactors, we performed streptavidin affinity purification (including RNase treatment) and immunoprecipitation mass spectrometry (IP-MS). Using tagged SKIV2L as bait, we identified various RNA binders, Ski complex components WDR61 and TTC37, and ribosomal proteins (Figure 3A; Table S5), consistent with the SKIV2L-ribosome interaction detected by CRAC (Figure 2A). To enrich for more direct SKIV2L interactions, we repeated the experiment adding a FLAG immunoprecipitation (tandem IP-MS). This recovered just two proteins, AVEN and FOCAD (KIAA1797), besides the Ski complex (Figure 3B).

FOCAD is a poorly characterized protein whose loss is associated with glioma (Brockschmidt et al., 2012) and colorectal cancer (Weren et al., 2015). Remarkably, its *Arabidopsis* homolog binds the Ski complex (Lange et al., 2019). AVEN is widely expressed and contributes to acute leukemia/lymphoma (Eißmann et al., 2013). Its disordered N-terminal glycine- and arginine-rich (RGG/RG) domain (Figure 3C) interacts with RNA

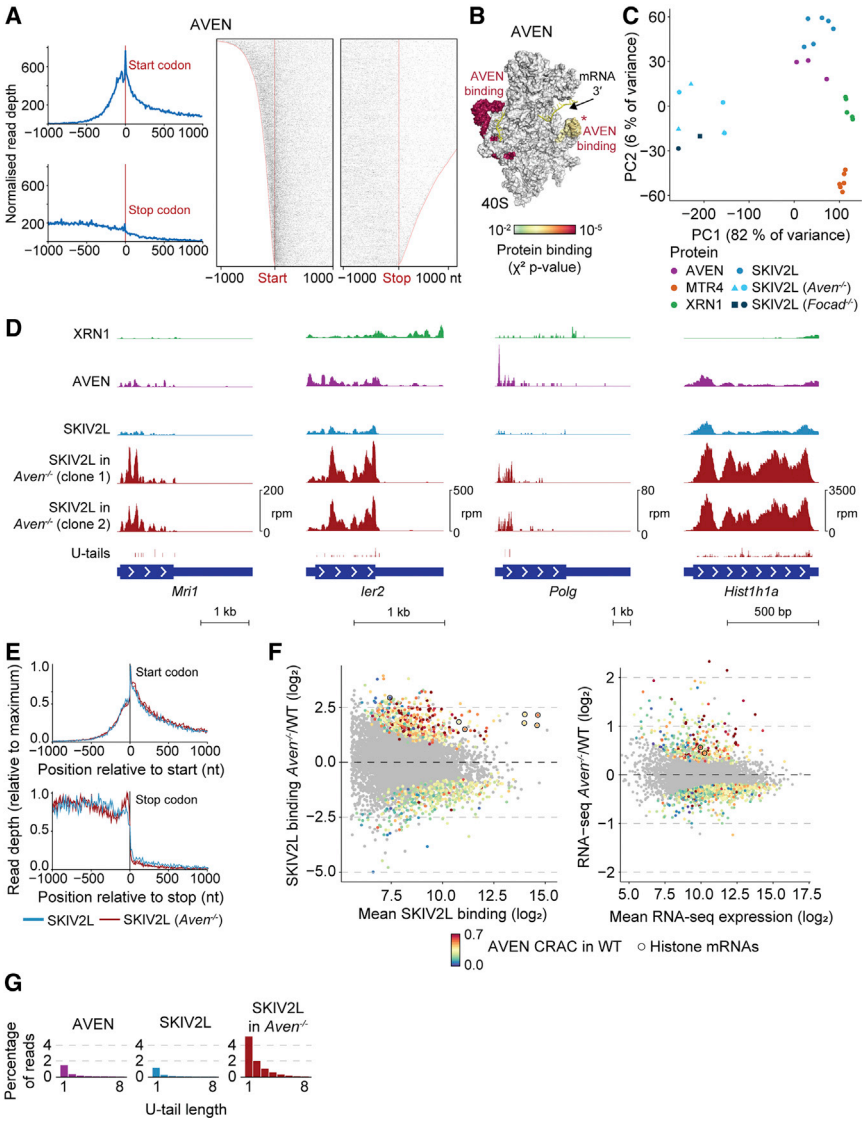



**Figure 4. SKIV2L Binding and 3′–5′ Decay Increase upon *Aven* Knockout**

(A and B) CRAC signal for AVEN around start and stop codons (A) and on the ribosomal 40S subunit (B).

(C) PCA based on mRNA counts. Shapes indicate different clones.

(D) CRAC coverage for individual mRNAs.

(E) SKIV2L CRAC around start and stop codons in WT and *Aven*$^{-/-}$ cells.

(F) Changes in SKIV2L CRAC binding (left) and RNA-seq counts (right) for *Aven*$^{-/-}$ versus WT cells. Significantly up/downregulated transcripts (padj < 0.05) are colored by AVEN CRAC counts in WT cells, relative to SKIV2L+XRN1+MTR4 counts, and replication-dependent histone mRNAs are circled.

(G) Proportion of AVEN or SKIV2L CRAC reads in mRNAs with 3′ U-tails.

See also Figure S3 and Table S3.

and localizes AVEN to polysomes (Thandapani et al., 2015). Furthermore, AVEN aids translation through G-quadruplexes in two mRNAs, and IP-MS using human AVEN as bait retrieved the Ski complex and FOCAD (Thandapani et al., 2015). These studies support our MS results.

To confirm the SKIV2L-AVEN interaction, we endogenously 3xFLAG-Avi-tagged *Aven* (Figure 3D) and performed IP-MS with AVEN as bait, recovering the Ski complex and FOCAD (Figure 3E). We repeated the tandem IP-MS using a higher salt concentration and SKIV2L or AVEN as bait (Figures 3F and 3G). AVEN now recovered FOCAD, but not the Ski complex, suggesting AVEN-FOCAD and SKIV2L-WDR61-TTC37 (Ski complex) are separable complexes that associate transiently with each other.

## SKIV2L Binding and 3′–5′ Decay Increase upon *Aven* Knockout

As AVEN associates with polysomes (Thandapani et al., 2015) and the Ski complex (Figure 3B), we speculated it might recruit SKIV2L to targets. To test this, we performed CRAC on 3xFLAG-Avi-tagged AVEN mESCs to map AVEN-binding sites. Like SKIV2L, AVEN bound the 5′ UTR and CDS of mRNAs (Figure 4A), albeit with a stronger 5′ bias. AVEN CRAC also revealed ribosome contacts, one overlapping that of SKIV2L (Figure 4B, marked with an asterisk [*]). PCA based on mRNA binding revealed that AVEN and SKIV2L bound common targets (Figures 4C and S3A), and AVEN and SKIV2L bound similar regions on individual mRNAs (Figure 4D). These similarities suggest that AVEN and SKIV2L function in the same pathway.

To determine whether AVEN affects Ski complex recruitment to mRNAs, we generated *Aven*$^{-/-}$ mESCs by deleting the C-terminal portion of the protein (Figure S3B). This led to near-complete knockdown of the entire *Aven* mRNA (Figure S3B). SKIV2L CRAC revealed that while the average binding pattern of SKIV2L along mRNAs was unaffected in *Aven*$^{-/-}$ (Figure 4E), there were strong differences in which mRNAs were bound, apparent from a PCA (Figures 4C and S3A). AVEN thus plays a role in SKIV2L targeting. SKIV2L binding was similarly perturbed in *Focad*$^{-/-}$ mESCs (Figures 4C and S3C), suggesting that AVEN and FOCAD functionally overlap. Due to its size and low abundance, FOCAD was challenging to work with, so we focused on AVEN.

In contrast to our prediction, SKIV2L binding to mRNAs was not reduced in *Aven*$^{-/-}$ cells but instead increased at many sites (Figure 4F; examples in Figure 4D). To account for changes in RNA abundance, we normalized CRAC to RNA-seq counts from WT and *Aven*$^{-/-}$ cells (Table S3). Increased SKIV2L binding was accompanied by elevated 3′ uridylation of bound RNAs (Figure 4G), indicating increased 3′–5′ decay. This suggests that

unlike WT conditions, where SKIV2L transiently scans all translation events, upon *Aven* deletion, SKIV2L assists intensively in 3′–5′ decay at specific sites. These sites are bound by AVEN in WT cells (Figure 4F, left, and Figure S3D), exemplified by replication-dependent histone mRNAs (circled in Figure 4F), suggesting that changes in SKIV2L binding are a direct consequence of losing AVEN. Changes in mRNA levels in *Aven*$^{-/-}$ cells (Figure 4F, right) were smaller than changes in SKIV2L CRAC and correlate poorly with AVEN binding (Figure S3E) so likely represent secondary effects.

In summary, AVEN does not recruit the Ski complex. Instead, loss of AVEN increases SKIV2L binding and 3′–5′ RNA decay at many sites. As aberrant translation events recruit SKIV2L and AVEN may assist translation (Thandapani et al., 2015), we hypothesize that AVEN prevents ribosome stalls that otherwise trigger SKIV2L binding and mRNA decay.

## AVEN and SKIV2L Counteract Ribosome Stalling

To globally assess how AVEN affects translation and ribosome stalling, we performed monosome and disome profiling for WT and *Aven*$^{-/-}$ mESCs (Table S4). We plotted changes in mRNA disome and monosome densities (Figure 5A), distinguishing mRNAs with increased, decreased, or unchanged SKIV2L binding in *Aven*$^{-/-}$ versus WT (pink/blue/gray points in Figure 5A) and calculated best fit lines. This revealed that in *Aven*$^{-/-}$, changes in monosome and disome density occur for all categories of mRNAs and are correlated, as expected. However, on top of these changes, mRNAs gaining SKIV2L binding in *Aven*$^{-/-}$ display a further increase in disome occupancy (upward shift of pink points in Figure 5A; exemplified by replication-dependent histone mRNAs in Figures 5B and 5C). mRNAs accumulating disomes upon *Aven* knockout were bound by AVEN in WT (Figure S4A), and disome changes in individual transcripts overlapped with AVEN and SKIV2L binding (Figure 5C). These data suggest that stalled ribosomes accumulating in *Aven*$^{-/-}$ drive increased SKIV2L recruitment, which presumably clears these mRNAs.

According to this model, the combined absence of AVEN and SKIV2L should have an additive effect, as SKIV2L would not be available to clear stalled messenger RNPs (mRNPs) arising in the absence of AVEN. AVEN targets should thus be stabilized and accumulate in a double knockout. To test this, we generated a *Skiv2l*$^{fl/fl}$ conditional knockout in *Aven*$^{-/-}$ mESCs and performed RNA-seq after 4OHT treatment (Figure S4B). In contrast to the single *Skiv2l*$^{fl/fl}$ knockout, where transcripts accumulated after 4 days of 4OHT (Figure 1I), we observed widespread changes in *Aven*$^{-/-}$ *Skiv2l*$^{fl/fl}$ after 2 days (Figure 5D; Table S3). Upregulated transcripts displayed high AVEN binding in WT (Figure 5D, left) and increased SKIV2L occupancy in *Aven*$^{-/-}$ (Figure 5D, right), suggesting they are direct SKIV2L and AVEN targets. Transcriptome-wide half-life measurements following actinomycin D transcription shut off confirmed that these targets are stabilized in the double knockout (Figure S4C; Table S3). The accumulation of replication-dependent histone mRNAs was particularly striking (Figures 5D, circled). These results support a model whereby AVEN and SKIV2L cooperate in translation-coupled RNA surveillance, with AVEN opposing translational stalls and SKIV2L eliminating mRNAs if aberrant events accumulate. Furthermore, SKIV2L and AVEN maintain normal histone translation and RNA levels.

## AVEN and SKIV2L Affect Expression of Many mRNAs

As replication-dependent histone levels are coupled to DNA synthesis, with histone mRNAs accumulating until they are degraded at the end of S-phase, we suspected that perturbed histone expression in the absence of AVEN and SKIV2L might alter cell-cycle progression. To test this, we synchronized mESCs at G1/S using a double thymidine block and monitored DNA content by DAPI staining following release (Figure S4D). *Aven*$^{-/-}$ *Skiv2l*$^{fl/fl}$ double knockout cells exhibited delayed progression through S phase, into G2, and ultimately into G1, in line with their altered histone mRNA abundance (compared to WT or single knockouts). We conclude that the AVEN-SKIV2L pathway contributes to cell cycle progression.

While examining individual mRNAs, we noticed that besides main CDS regions, SKIV2L and ribosomes also accumulate in uORFs in *Aven*$^{-/-}$ (Figures 5E and 5F). AVEN occupied these uORFs in WT cells (Figure 5F), and increased ribosome occupancy in *Aven*$^{-/-}$ cells correlated with WT AVEN binding (Figure S4E) and increased SKIV2L binding in *Aven*$^{-/-}$ (Figure 5E). Whereas *Aven* knockout resulted in increased disome occupancy in main CDS regions, changes across uORFs occurred for monosomes, disomes, or both. AVEN thus has a complex effect on 5′ UTR translation.

As uORF translation can alter mRNA stability or main CDS translation (Calvo et al., 2009), we wondered whether such changes occur upon loss of AVEN and/or SKIV2L. We focused on *Atf4* and *Ifrd1* mRNAs, with functional uORFs bound by AVEN and SKIV2L (Figure 5F). Under normal conditions, *Ifrd1* uORF translation destabilizes the mRNA via NMD (Zhao et al., 2010). *Ifrd1* RNA accumulated after 4 days of *Skiv2l* knockout and 2 days of *Aven Skiv2l* double knockout (Figure 5G), suggesting that SKIV2L participates in *Ifrd1* mRNA clearance, and this is enhanced by increased uORF ribosome occupancy in *Aven*$^{-/-}$.

In contrast to the destabilizing effect of the *Ifrd1* uORF, within *Atf4*, two uORFs modulate translation of the main CDS (Harding et al., 2000; Vattem and Wek, 2004). Ribosomes normally translate uORF1 then reinitiate at uORF2, preventing them from translating the main CDS, but during the integrated stress response (ISR), phosphorylation of the translation factor eIF2$\alpha$ reduces preinitiation complex availability. Ribosomes now scan past uORF2 and reinitiate at the downstream main CDS, producing ATF4 protein. To examine the effects of increased ribosome occupancy over *Atf4* uORFs in *Aven*$^{-/-}$, we monitored ATF4 accumulation upon activation of the ISR with thapsigargin. Compared to WT cells, ATF4 levels peaked earlier in *Aven*$^{-/-}$, despite similar levels of eIF2$\alpha$ phosphorylation and basal ATF4 pre-induction (Figure S4F). Therefore, binding of AVEN and SKIV2L to uORFs modifies transcript stability (*Ifrd1*) and main CDS translation (*Atf4*).

In summary, the roles played by AVEN and SKIV2L in counteracting aberrant translation are crucial for expression of uORF-containing and histone mRNAs, among others.

## AVEN Acts on Structured RNAs

We next asked what makes mRNAs dependent on AVEN. AVEN crosslinks to G-quadruplexes in *Mll1* and *Mll4* mRNAs

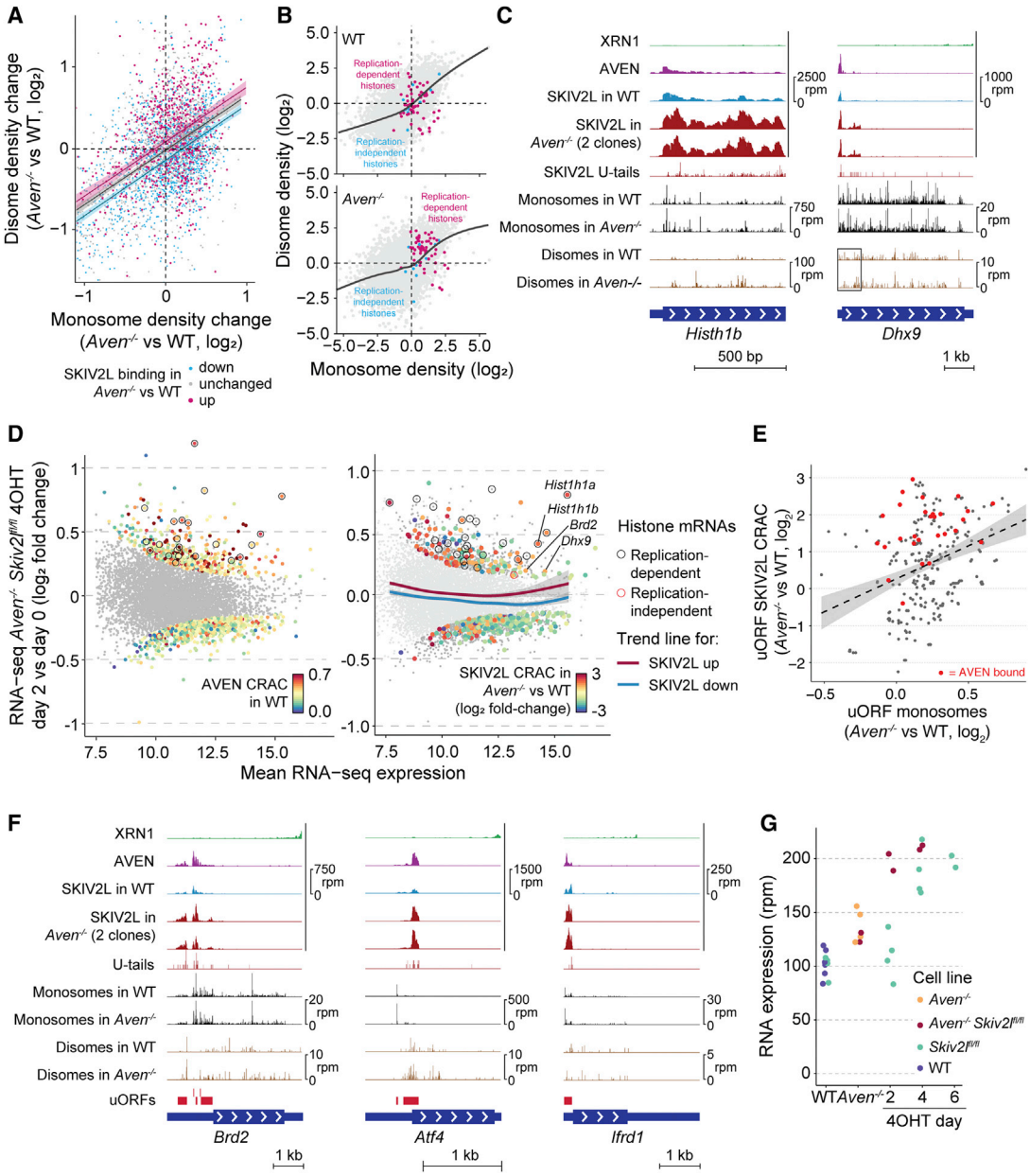



**Figure 5. AVEN and SKIV2L Counteract Ribosome Stalling**

(A) Changes in mRNA monosome and disome densities in *Aven*$^{-/-}$ versus WT. Transcripts are colored by changes in SKIV2L binding in *Aven*$^{-/-}$ versus WT (threshold $\log_2$ fold change = ±0.5; up, n = 1,856; down, n = 2,019; unchanged, n = 2,373), and a linear best fit is plotted for each group (shaded area represents 95% confidence interval).

(B) Monosome and disome densities in WT (top) and *Aven*$^{-/-}$ (bottom), highlighting histone mRNAs and with a cubic regression trendline.

(C) CRAC and monosome/disome profiling for individual mRNAs.

(D) Changes in mRNA abundance for *Skiv2l*$^{fl/fl}$ *Aven*$^{-/-}$ cells after 2-day 4-OHT. Significantly changing mRNAs (padj < 0.05) colored by AVEN CRAC in WT (left) or SKIV2L CRAC changes in *Aven*$^{-/-}$ versus WT (right). Cubic regression trendlines are shown for all mRNAs, grouped by increased/decreased SKIV2L binding in *Aven*$^{-/-}$ versus WT (right).

(E) Changes in uORF SKIV2L CRAC and monosome profiling counts for *Aven*$^{-/-}$ versus WT cells. Both datasets normalized to main CDS monosome profiling counts. A best-fit line is shown, with 95% confidence intervals, and AVEN-bound uORFs (defined in Figure S4E) are colored red.

(F) CRAC and monosome/disome profiling for individual mRNAs. uORFs identified from monosome profiling are shown in red.

(G) RNA-seq counts for *Ifrd1* in various cell lines showing individual replicates.

See also Figure S4 and Tables S3 and S4.

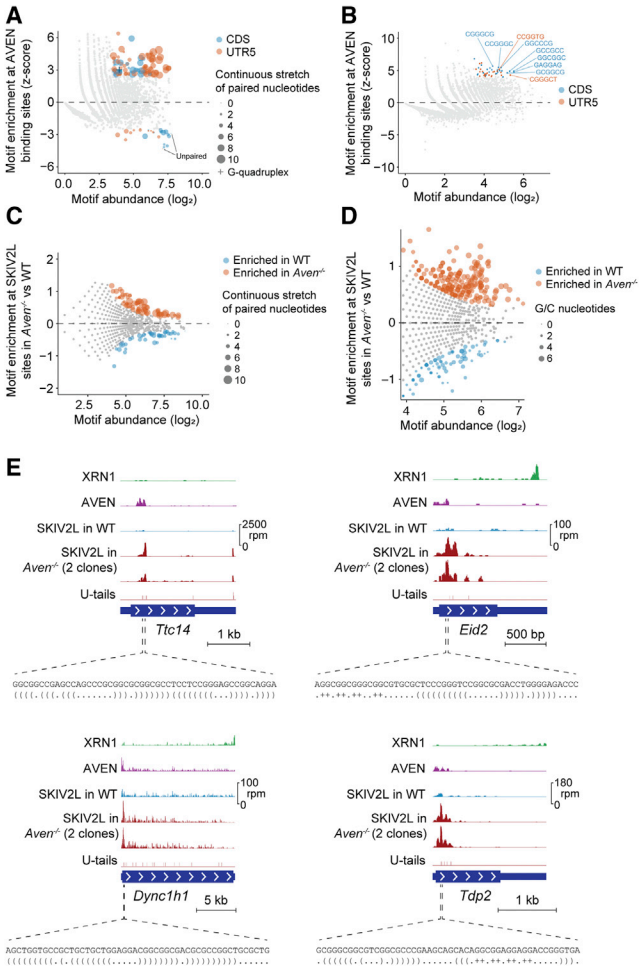



**Figure 6. AVEN Acts on Structured RNAs**

(A and B) Structure (A) and sequence motif analysis (B) for AVEN-binding sites, based on CRAC versus RNA-seq enrichments in 50-nt 5′ UTR and CDS windows. Points (structure motifs) in (A) are scaled by paired nucleotide content. (C and D) Structure (C) and sequence motif analysis (D) for SKIV2L binding sites, comparing *Aven*$^{-/-}$ and WT cells. Points in (C) are scaled as for (A), and points in (D) are scaled by GC content.

(E) Examples of AVEN-bound windows showing their sequence, predicted structure (bracket/dot annotation for paired/unpaired nucleotides; + = G-quadruplex), and CRAC coverage for various proteins.

See also Figure S5.

(Thandapani et al., 2015), and RGG/RG motifs like AVEN's can melt G-rich or G-quadruplex sequences (Loughlin et al., 2019; Meyer et al., 2019). To test whether AVEN binds specific RNA sequences or structures, we examined the highest AVEN-bound 50-nt windows from each mRNA 5′ UTR and CDS (based on CRAC). Compared to control regions, AVEN-bound regions were enriched for stretches of paired nucleotides or G-quadruplexes, based on RNAfold predictions (Figure 6A), and GC-rich sequences (Figure 6B). The same was true of regions binding SKIV2L in *Aven*$^{-/-}$ (Figures 6C and 6D), and these preferences were clear for individual mRNAs (Figure 6E). This suggests that AVEN binding to GC-rich sites with structural propensity avoids sustained SKIV2L recruitment.

Interestingly, many SKIV2L-bound RNA fragments possessed 3′ U-tails in *Aven*$^{-/-}$ cells (Figure 4G), particularly where SKIV2L binding increased (Figure S5A). As U-tails are added to 5′ RNA cleavage products, we reasoned they could pinpoint sites of mRNA cleavage and decay in *Aven*$^{-/-}$. Indeed, U-tailed SKIV2L-bound RNA fragments in *Aven*$^{-/-}$ were enriched upstream of predicted structured regions (Figure S5B). Disomes aligned here in *Aven*$^{-/-}$ (but not WT), and SKIV2L binding increased (Figure S5B). This suggests that structure-prone regions impede translation in *Aven*$^{-/-}$, leading to ribosome stalling, RNA cleavage, SKIV2L recruitment, and decay. We speculate that AVEN helps suppress or melt RNA structures, consistent with its binding to structure-prone regions.

## sORF Surveillance by AVEN and SKIV2L

As SKIV2L and AVEN specialize in translation surveillance, we did not expect them to bind non-coding RNAs (ncRNAs). However, upon *Aven* knockout, SKIV2L bound transcripts from intergenic, upstream, and antisense loci, and AVEN bound these ncRNAs in WT cells (see examples in Figure 7A). To examine this globally we divided the genome into 1 kb windows classified by protein-coding gene overlap, and calculated $\log_2$-fold changes in SKIV2L CRAC in *Aven*$^{-/-}$ versus WT (Table S6). This revealed increased SKIV2L binding to RNAs from hundreds of non-coding regions (Figure 7B), accompanied by increased U-tailing (Figure S5A). These transcripts were GC rich and predicted to form strong secondary structures (Figures S6A and S6B), suggesting the same mechanism drives SKIV2L recruitment to ncRNAs and mRNAs upon *Aven* deletion.

We wondered whether SKIV2L binding is due to ectopic ribosome occupancy on these "non-coding" RNAs. Indeed, monosomes and disomes accumulated at sites bound by SKIV2L in *Aven*$^{-/-}$ (e.g., Figure 7A), which often overlapped small ORFs (sORFs; Table S7), suggesting they are translated. Looking globally, we calculated $\log_2$-fold changes in monosome and disome counts for the non-coding 1-kb windows defined in Figure 7B, classified by differential SKIV2L binding in *Aven*$^{-/-}$. This revealed a correlation between gain of SKIV2L binding and increased monosome and disome occupancy (Figure 7C, "SKIV2L CRAC," "Monosomes", and "Disomes"). Elevated disome occupancy was particularly strong, suggesting increased ribosome stalling. The peptides generated by these translation events do not appear to perform conserved functions, as their sequences have low phyloCSF scores (Figure S6C). Changes in SKIV2L binding correlated with AVEN occupancy in WT cells, supporting a direct role for AVEN (Figure 7C, "AVEN vs SKIV2L CRAC"). Overall, our data suggest that loss of AVEN results in ribosome stalling on sORFs in structured ncRNAs, which is resolved by surveillance involving RNA cleavage and SKIV2L-dependent decay.

A prediction of this is that upon *Aven* deletion, these ncRNAs should become reliant on SKIV2L-dependent 3′–5′ decay, which specializes in degrading RNAs with arrested ribosomes. Presumably, alternative pathways remove these transcripts when AVEN is present. Indeed, these ncRNAs do not strongly accumulate in *Aven*$^{-/-}$ and only slowly accumulate upon *Skiv2l*

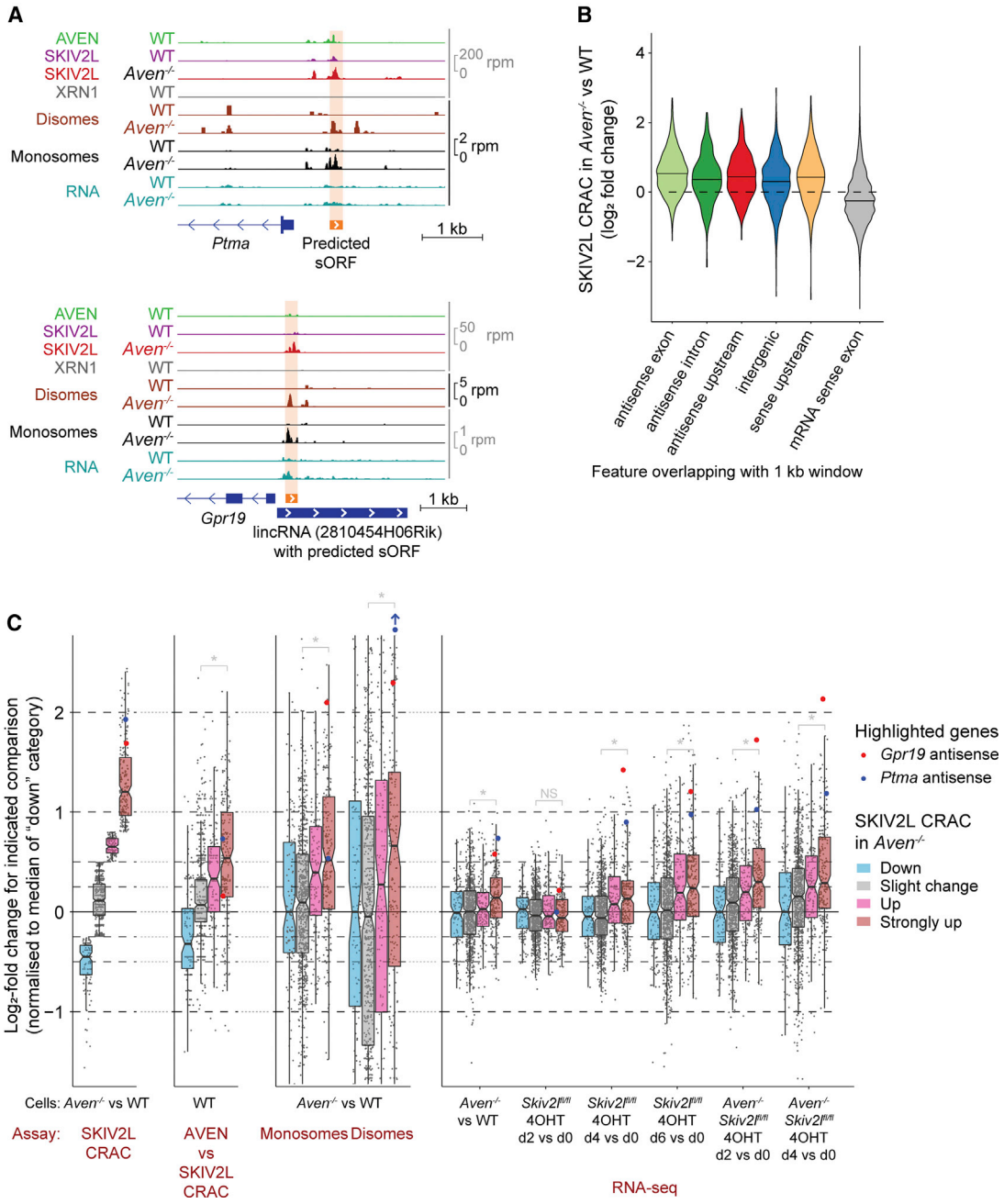



**Figure 7. Small ORF Surveillance by AVEN and SKIV2L**

(A) CRAC, monosome, and disome profiling across ncRNA regions in WT and *Aven*$^{-/-}$ cells, with small ORFs indicated. RNA-seq of ribosome profiling inputs shown in blue.

(B) SKIV2L CRAC changes (*Aven*$^{-/-}$ versus WT) for 1-kb genomic windows classified by overlap with protein-coding genes.

(C) CRAC, ribosome profiling, and RNA-seq changes for the indicated comparisons, for non-coding 1-kb genomic windows defined in (B). Windows are categorized by their change in SKIV2L CRAC for *Aven*$^{-/-}$ versus WT cells (defined in the leftmost plot). The two genes in (A) are highlighted. *$p < 10^{-4}$ ("Slight change" versus "Strongly up" categories; Mann-Whitney U test with Bonferroni correction).

See also Figure S6 and Tables S6 and S7.

knockout but rapidly accumulate when *Skiv2l* is knocked out in *Aven*$^{-/-}$ cells (Figure 7C, "RNA-seq"). We conclude that the absence of AVEN renders cells dependent on SKIV2L to clear ncRNAs with trapped ribosomes. The AVEN-SKIV2L pathway thus plays a universal role in counteracting aberrant translation on coding RNAs and ncRNAs.

## DISCUSSION

### Mammalian mRNA Decay: Specialization and Links to Translation

We are struck by the widespread coupling between cytoplasmic mRNA decay and translation revealed by our study. Evidence of such crosstalk has been mounting, from reports that SKIV2L and XRN1 associate with polysomes (Mangus and Jacobson, 1999; Qu et al., 1998) to analyses of decay intermediates (Antic et al., 2015; Hu et al., 2009; Pelechano et al., 2015) and structures of the Ski complex and Xrn1 bound to yeast ribosomes (Schmidt et al., 2016; Tesina et al., 2019). We show that XRN1 and SKIV2L ribosome binding sites are conserved to mammals, these interactions occur under physiological conditions, and remarkably, SKIV2L is exclusively and universally recruited by ribosomes.

Ski2 was thought to act redundantly with Xrn1 in bulk RNA decay, based on synthetic lethality in yeast (Anderson and Parker, 1998; Johnson and Kolodner, 1995). However, yeast Ski2 binding to 3′ UTRs (Sohrabi-Jahromi et al., 2019; Tuck and Tollervey, 2013) relies on fungus-specific factors such as Ska1 to antagonize ribosome interactions (Zhang et al., 2019). Our data argue that mammalian SKIV2L does not function in full-length mRNA decay but acts almost exclusively in translation-associated RNA surveillance. As the Ski complex is indispensable for cytoplasmic exosome activity (Anderson and Parker, 1998; Araki et al., 2001; van Hoof et al., 2000), this implies that the cytoplasmic exosome acts similarly exclusively in surveillance. We note that mammals possess an exosome-independent 3′–5′ decay pathway (DIS3L2). This might assist XRN1 in bulk decay, in line with a report that XRN1 and DIS3L2 knockdowns result in broader mRNA changes than DIS3L (exosome) knockdown (Lubas et al., 2013). The 3′ UTR accumulation of XRN1 suggests a passive role for translation in 5′–3′ decay. Future biochemical studies should help clarify these possible differences between SKIV2L- and XRN1-ribosome interactions.

Interestingly, we found that SKIV2L acts in bulk decay of a few mRNAs. Unique features might render these accessible to, or dependent on, ribosome-bound SKIV2L. For example, cleavage of *Ifrd1* might generate an access point for SKIV2L (Ottens et al., 2017), and ribosome-bound SKIV2L could reach the end of short histone mRNA 3′ UTRs. This was proposed for *S. cerevisiae* mRNAs (Zhang et al., 2019), and we see SKIV2L binding to very short 3′ UTRs (Figure 2B). Alternatively, surveillance-inducing ribosome collisions may be rife within histone mRNAs, whose decay requires stalled ribosome factors HBS1 and PELOTA (Slevin et al., 2014). Although this pathway is wasteful, as it eliminates the nascent polypeptide, for replication-dependent histones, this may help to tightly restrict their expression to S-phase.

For most mRNAs, however, there is a clear division of labor, with XRN1 specializing in bulk RNA decay (albeit with a minor role in surveillance) and SKIV2L in surveillance. This ensures that translation is not interrupted by bulk RNA turnover, as XRN1 follows the last ribosome, and may reflect a need for dedicated surveillance factors to wrestle mRNAs from arrested ribosomes. Indeed, it is even possible that SKIV2L could perform additional roles in resolving stalled mRNA-ribosome complexes, beyond assisting the exosome in 3′–5′ decay.

### Defining Translation-Dependent mRNA Surveillance

Our data also reveal triggers and components of RNA surveillance. SKIV2L pervasively interacts with ribosome-occupied regions, establishing it as a central component of translation surveillance. Based on the low level of U-tailing (a proxy for RNA cleavage), we suggest SKIV2L binding in WT cells mostly reflects dynamic probing of translation, which rarely triggers a full surveillance response. Nonetheless, SKIV2L and disomes were enriched at A-rich tracts, proline sequences, and uORFs, suggesting they occasionally trigger ribosome stalling and RNA decay. For A-rich tracts, the sequence appears key, consistent with reports that ~11 As attenuate translation in human cells (Arthur et al., 2015). We find this occurs at many endogenous sites with 8–9 As sufficient.

Besides defining SKIV2L targets, we established AVEN and FOCAD as components of this pathway. AVEN was reported to interact with the Ski complex and FOCAD in human cells (Thandapani et al., 2015) and identified in an NMD screen (Alexandrov et al., 2017), and the plant FOCAD homolog Rst1 interacts with the Ski complex and exosome (Lange et al., 2019). AVEN is conserved from mammals to flies (Zou et al., 2011) and FOCAD to plants (Lange et al., 2019), so their RNA decay roles may be evolutionarily important.

### AVEN as an Anti-stalling Factor

We propose that AVEN prevents ribosome stalls, which otherwise trigger mRNA cleavage and decay. The RNA-binding preferences and position of AVEN on the ribosome might let it directly melt structures arresting translation, potentially via its RGG/RG domain. Supporting this, FUS and AUF1 RGG/RG domains remodel RNA (Loughlin et al., 2019; Meyer et al., 2019). Alternatively, AVEN might recruit a helicase (Thandapani et al., 2015), although besides SKIV2L, we did not detect helicase partners for AVEN.

In our model, AVEN acts prior to SKIV2L, to prevent ribosome stalling, and is potentially loaded with scanning ribosomes. However, our IP-MS data suggest that AVEN and SKIV2L directly interact. To resolve this paradox, we propose that the AVEN-SKIV2L interaction is transient, perhaps serving as a handover to ensure unresolved ribosome stalls are not left unchecked. Transient "connections" are common in RNA surveillance, as reported for Ski complex-exosome (Kalisiak et al., 2017) and nuclear MTR4-ZFC3H1-PABPN1 interactions (Meola et al., 2016).

Exploring the AVEN-SKIV2L pathway revealed that uORF-containing and histone mRNAs are particularly sensitive. AVEN prevents cell-cycle arrest in osteosarcoma and *Drosophila* cells (Baranski et al., 2015; Zou et al., 2011) and delays mitotic entry in *Xenopus* egg extracts (Guo et al., 2008; Zou et al., 2011). Our data suggest AVEN also plays a direct role in cell-cycle progression via reducing ribosome stalling on histone mRNAs. The most surprising AVEN and SKIV2L substrates, however, were ncRNAs. Here, an appealing model is that AVEN assists in functional small peptide production. Although AVEN-dependent sORFs have low phyloCSF scores and we

could not detect derived peptides, AVEN could enable cells to express peptides that eventually evolve to become stable and perform important roles. Alternatively, AVEN and SKIV2L may target nuclear ncRNAs escaping to the cytoplasm. These structured RNAs could function in the nucleus but in the cytoplasm might become stuck on ribosomes if left unchecked.

In conclusion, we find that mammalian RNA decay pathways are highly specialized and cytoplasmic decay is widely coupled to translation. While normal translation may assist bulk mRNA turnover, aberrant translation events pose a diverse threat counteracted by the concerted activity of AVEN and SKIV2L.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Generation of endogenously tagged cell lines
  - Generation of straight KO cell lines
  - Generation of conditional KO cell lines
  - Transfections
  - RNA sequencing
  - CRAC
  - Translation inhibition experiments for CRAC
  - Ribosome profiling
  - Western Blotting
  - Co-immunoprecipitations
  - Affinity purification for LC–MS/MS
  - Mass spectrometry analysis
  - Cell cycle analysis
  - ATF4 induction with Thapsigargin
  - RT-qPCR
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - CRAC data preprocessing and alignment
  - CRAC quantification of non-templated 3′ tails
  - CRAC PCA, correlation matrix and tSNE
  - Differential SKIV2L binding analysis
  - Identification of rRNA binding sites by CRAC
  - CRAC plots around start and stop codons
  - CRAC enrichment at amino acid combinations
  - CRAC and disome profiling repeat analyis
  - RNA-seq analysis
  - RNA half-life analysis
  - CRAC sequence and structure motif analysis
  - CRAC/ribosome profiling at structured regions
  - CRAC/ribosome profiling for genomic windows
  - Ribosome profiling analysis
- DATA AND CODE AVAILABILITY

## SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2020.01.007.

## ACKNOWLEDGMENTS

We thank the FMI Functional Genomics, FACS and Protein Analysis facilities, and the University of Lausanne Genomics Technology Facility. This work was supported by the Swiss National Science Foundation (SNSF), NCCR RNA & Disease (grant 141735), and the Novartis Research Foundation. A.C.T. is supported by the Wellcome Trust (103977), and D.G. is supported by the SNSF (grant 179190) and the University of Lausanne.

## AUTHOR CONTRIBUTIONS

A.C.T. and A.R. conceived the project, designed and performed experiments, analyzed data, and wrote the manuscript. A.C.T. performed bioinformatics analyses. M.B. and D.G. conceived the project, designed experiments, and edited the manuscript. A.B.A. and V.C.-S. analyzed ribosome profiling data. L.A.L. performed ribosome profiling experiments. D.H. and V.I. performed MS analyses.

## DECLARATION OF INTERESTS

The Friedrich Miescher Institute for Biomedical Research (FMI) receives significant financial contributions from the Novartis Research Foundation.

Received: July 29, 2019
Revised: November 11, 2019
Accepted: January 7, 2020
Published: February 10, 2020

## REFERENCES

Alexandrov, A., Shu, M.-D., and Steitz, J.A. (2017). Fluorescence amplification method for forward genetic discovery of factors in human mRNA degradation. Mol. Cell *65*, 191–201.

Anderson, J.S., and Parker, R.P. (1998). The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. EMBO J. *17*, 1497–1506.

Antic, S., Wolfinger, M.T., Skucha, A., Hosiner, S., and Dorner, S. (2015). General and microRNA-mediated mRNA degradation occurs on ribosome complexes in Drosophila cells. Mol. Cell. Biol. *35*, 2309–2320.

Araki, Y., Takahashi, S., Kobayashi, T., Kajiho, H., Hoshino, S., and Katada, T. (2001). Ski7p G protein interacts with the exosome and the Ski complex for 3′-to-5′ mRNA decay in yeast. EMBO J. *20*, 4684–4693.

Arpat, A.B., Liechti, A., Matos, M.D., Dreos, R., Janich, P., and Gatfield, D. (2019). Transcriptome-wide sites of collided ribosomes reveal principles of translational pausing. bioRxiv. https://doi.org/10.1101/710061.

Arthur, L., Pavlovic-Djuranovic, S., Smith-Koutmou, K., Green, R., Szczesny, P., and Djuranovic, S. (2015). Translational control by lysine-encoding A-rich sequences. Sci. Adv. *1*, e1500154.

Baranski, Z., Booij, T.H., Cleton-Jansen, A.M., Price, L.S., van de Water, B., Bovée, J.V., Hogendoorn, P.C., and Danen, E.H. (2015). Aven-mediated checkpoint kinase control regulates proliferation and resistance to chemotherapy in conventional osteosarcoma. J. Pathol. *236*, 348–359.

Bresson, S., Tuck, A., Staneva, D., and Tollervey, D. (2017). Nuclear RNA decay pathways aid rapid remodeling of gene expression in yeast. Mol. Cell *65*, 787–800.

Brockschmidt, A., Trost, D., Peterziel, H., Zimmermann, K., Ehrler, M., Grassmann, H., Pfenning, P.N., Waha, A., Wohlleber, D., Brockschmidt, F.F., et al. (2012). KIAA1797/FOCAD encodes a novel focal adhesion protein with tumour suppressor function in gliomas. Brain *135*, 1027–1041.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. USA *106*, 7507–7512.

Castelo-Szekely, V., De Matos, M., Tusup, M., Pascolo, S., Ule, J., and Gatfield, D. (2019). Charting DENR-dependent translation reinitiation uncovers

predictive uORF features and links to circadian timekeeping via Clock. Nucleic Acids Res. *47*, 5193–5209.

Chothani, S.P., Adami, E., Viswanathan, S., Hubner, N., Cook, S., Schafer, S., and Rackham, O.J.L. (2017). Reliable detection of translational regulation with Ribo-seq. bioRxiv. https://doi.org/10.1101/234344.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

D'Orazio, K.N., Wu, C.C.-C., Sinha, N., Loll-Krippleber, R., Brown, G.W., and Green, R. (2019). The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during No Go Decay. eLife *8*, e49117.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Eißmann, M., Melzer, I.M., Fernández, S.B.M., Michel, G., Hrabě de Angelis, M., Hoefler, G., Finkenwirth, P., Jauch, A., Schoell, B., Grez, M., et al. (2013). Overexpression of the anti-apoptotic protein AVEN contributes to increased malignancy in hematopoietic neoplasms. Oncogene *32*, 2586–2591.

Fabre, A., Charroux, B., Martinez-Vinson, C., Roquelaure, B., Odul, E., Sayar, E., Smith, H., Colomb, V., Andre, N., Hugot, J.-P., et al. (2012). SKIV2L mutations cause syndromic diarrhea, or trichohepatoenteric syndrome. Am. J. Hum. Genet. *90*, 689–692.

Flemr, M., and Bühler, M. (2015). Single-step generation of conditional knockout mouse embryonic stem cells. Cell Rep. *12*, 709–716.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2013). Ensembl 2013. Nucleic Acids Res. *41*, D48–D55.

Gatfield, D., and Izaurralde, E. (2004). Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature *429*, 575–578.

Ghosh, S., and Jacobson, A. (2010). RNA decay modulates gene expression and controls its fidelity. Wiley Interdiscip. Rev. RNA *1*, 351–361.

Glover, M.L., Burroughs, A.M., Egelhofer, T.A., Pule, M.N., Aravind, L., and Arribere, J.A. (2019). NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. bioRxiv. https://doi.org/10.1101/674358.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proc. Natl. Acad. Sci. USA *106*, 9613–9618.

Grozdanov, P., Georgiev, O., and Karagyozov, L. (2003). Complete sequence of the 45-kb mouse ribosomal DNA repeat: analysis of the intergenic spacer. Genomics *82*, 637–643.

Guo, J.Y., Yamada, A., Kajino, T., Wu, J.Q., Tang, W., Freel, C.D., Feng, J., Chau, B.N., Wang, M.Z., Margolis, S.S., et al. (2008). Aven-dependent activation of ATM following DNA damage. Curr. Biol. *18*, 933–942.

Guydosh, N.R., and Green, R. (2014). Dom34 rescues ribosomes in 3′ untranslated regions. Cell *156*, 950–962.

Guydosh, N.R., and Green, R. (2017). Translation of poly(A) tails leads to precise mRNA cleavage. RNA *23*, 749–761.

Halbach, F., Reichelt, P., Rode, M., and Conti, E. (2013). The yeast ski complex: crystal structure and RNA channeling to the exosome complex. Cell *154*, 814–826.

Harding, H.P., Novoa, I., Zhang, Y., Zeng, H., Wek, R., Schapira, M., and Ron, D. (2000). Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol. Cell *6*, 1099–1108.

Hartley, J.L., Zachos, N.C., Dawood, B., Donowitz, M., Forman, J., Pollitt, R.J., Morgan, N.V., Tee, L., Gissen, P., Kahr, W.H., et al. (2010). Mutations in TTC37 cause trichohepatoenteric syndrome (phenotypic diarrhea of infancy). Gastroenterology *138*, 2388–2398, 2398.e1–2398.e2.

Hsu, C.L., and Stevens, A. (1993). Yeast cells lacking 5′–>3′ exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5′ cap structure. Mol. Cell. Biol. *13*, 4826–4835.

Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., and Coller, J. (2009). Co-translational mRNA decay in Saccharomyces cerevisiae. Nature *461*, 225–229.

Ibrahim, F., Maragkakis, M., Alexiou, P., and Mourelatos, Z. (2018). Ribothrypsis, a novel process of canonical mRNA decay, mediates ribosome-phased mRNA endonucleolysis. Nat. Struct. Mol. Biol. *25*, 302–310.

Ikeuchi, K., Tesina, P., Matsuo, Y., Sugiyama, T., Cheng, J., Saeki, Y., Tanaka, K., Becker, T., Beckmann, R., and Inada, T. (2019). Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. EMBO J. *38*, e100276.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789–802.

Janich, P., Arpat, A.B., Castelo-Szekely, V., Lopes, M., and Gatfield, D. (2015). Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames. Genome Res. *25*, 1848–1859.

Johnson, A.W., and Kolodner, R.D. (1995). Synthetic lethality of sep1 (xrn1) ski2 and sep1 (xrn1) ski3 mutants of Saccharomyces cerevisiae is independent of killer virus and suggests a general role for these genes in translation control. Mol. Cell. Biol. *15*, 2719–2727.

Juszkiewicz, S., Chandrasekaran, V., Lin, Z., Kraatz, S., Ramakrishnan, V., and Hegde, R.S. (2018). ZNF598 is a quality control sensor of collided ribosomes. Mol. Cell *72*, 469–481.

Kalisiak, K., Kuliński, T.M., Tomecki, R., Cysewski, D., Pietras, Z., Chlebowski, A., Kowalska, K., and Dziembowski, A. (2017). A short splicing isoform of HBS1L links the cytoplasmic exosome and SKI complexes in humans. Nucleic Acids Res. *45*, 2068–2080.

Kearse, M.G., Goldman, D.H., Choi, J., Nwaezeapu, C., Liang, D., Green, K.M., Goldstrohm, A.C., Todd, P.K., Green, R., and Wilusz, J.E. (2019). Ribosome queuing enables non-AUG translation to be resistant to multiple protein synthesis inhibitors. Genes Dev. *33*, 871–885.

Khatter, H., Myasnikov, A.G., Natchiar, S.K., and Klaholz, B.P. (2015). Structure of the human 80S ribosome. Nature *520*, 640–645.

Kilchert, C., Wittmann, S., and Vasiljeva, L. (2016). The regulation and functions of the nuclear RNA exosome complex. Nat. Rev. Mol. Cell Biol. *17*, 227–239.

Knuckles, P., Carl, S.H., Musheev, M., Niehrs, C., Wenger, A., and Bühler, M. (2017). RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat. Struct. Mol. Biol. *24*, 561–569.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways: enzymes and mechanisms. Biochim. Biophys. Acta *1863*, 3125–3147.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713–724.

Lange, H., Ndecky, S.Y.A., Gomez-Diaz, C., Pflieger, D., Butel, N., Zumsteg, J., Kuhn, L., Piermaria, C., Chicher, J., Christie, M., et al. (2019). RST1 and RIPR connect the cytosolic RNA exosome to the Ski complex in Arabidopsis. Nat. Commun. *10*, 3871.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lim, J., Ha, M., Chang, H., Kwon, S.C., Simanshu, D.K., Patel, D.J., and Kim, V.N. (2014). Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell *159*, 1365–1376.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. *6*, 26.

Loughlin, F.E., Lukavsky, P.J., Kazeeva, T., Reber, S., Hock, E.-M., Colombo, M., Von Schroetter, C., Pauli, P., Cléry, A., Mühlemann, O., et al. (2019). The

solution structure of FUS bound to RNA reveals a bipartite mode of RNA recognition with both sequence and shape specificity. Mol. Cell *73*, 490–504.e6.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol. Cell *43*, 624–637.

Lubas, M., Damgaard, C.K., Tomecki, R., Cysewski, D., Jensen, T.H., and Dziembowski, A. (2013). Exonuclease hDIS3L2 specifies an exosome-independent 3′-5′ degradation pathway of human cytoplasmic mRNA. EMBO J. *32*, 1855–1868.

Mangus, D.A., and Jacobson, A. (1999). Linking mRNA turnover and translation: assessing the polyribosomal association of mRNA decay factors and degradative intermediates. Methods *17*, 28–37.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.

Meola, N., Domanski, M., Karadoulama, E., Chen, Y., Gentil, C., Pultz, D., Vitting-Seerup, K., Lykke-Andersen, S., Andersen, J.S., Sandelin, A., and Jensen, T.H. (2016). Identification of a nuclear exosome decay pathway for processed transcripts. Mol. Cell *64*, 520–533.

Meyer, A., Golbik, R.P., Sänger, L., Schmidt, T., Behrens, S.E., and Friedrich, S. (2019). The RGG/RG motif of AUF1 isoform p45 is a key modulator of the protein's RNA chaperone and RNA annealing activities. RNA Biol. *16*, 960–971.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′–>5′ exoribonucleases. Cell *91*, 457–466.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schübeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol. Cell *30*, 755–766.

Morton, D.J., Kuiper, E.G., Jones, S.K., Leung, S.W., Corbett, A.H., and Fasken, M.B. (2018). The RNA exosome and RNA exosome-linked disease. RNA *24*, 127–142.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev. *22*, 50–65.

Ostapcuk, V., Mohn, F., Carl, S.H., Basters, A., Hess, D., Iesmantavicius, V., Lampersberger, L., Flemr, M., Pandey, A., Thomä, N.H., et al. (2018). Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature *557*, 739–743.

Ottens, F., Boehm, V., Sibley, C.R., Ule, J., and Gehring, N.H. (2017). Transcript-specific characteristics determine the contribution of endo- and exonucleolytic decay pathways during the degradation of nonsense-mediated decay substrates. RNA *23*, 1224–1236.

Parker, R. (2012). RNA degradation in Saccharomyces cerevisae. Genetics *191*, 671–702.

Pashler, A.L., Towler, B.P., Jones, C.I., and Newbury, S.F. (2016). The roles of the exoribonucleases DIS3L2 and XRN1 in human disease. Biochem. Soc. Trans. *44*, 1377–1384.

Pavlov, M.Y., Watts, R.E., Tan, Z., Cornish, V.W., Ehrenberg, M., and Forster, A.C. (2009). Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proc. Natl. Acad. Sci. USA *106*, 50–54.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2015). Widespread co-translational RNA decay reveals ribosome dynamics. Cell *161*, 1400–1412.

Pérez-Ortín, J.E., Alepuz, P., Chávez, S., and Choder, M. (2013). Eukaryotic mRNA decay: methodologies, pathways, and links to other stages of gene expression. J. Mol. Biol. *425*, 3750–3775.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295.

Qu, X., Yang, Z., Zhang, S., Shen, L., Dangel, A.W., Hughes, J.H., Redman, K.L., Wu, L.-C., and Yu, C.Y. (1998). The human DEVH-box protein Ski2w from the HLA is localized in nucleoli and ribosomes. Nucleic Acids Res. *26*, 4068–4077.

Quinlan, A.R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. Curr. Protoc. Bioinformatics *47*, 1–34.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140.

Robinson, S.R., Oliver, A.W., Chevassut, T.J., and Newbury, S.F. (2015). The 3′ to 5′ exoribonuclease DIS3: from structure and mechanisms to biological functions and role in human disease. Biomolecules *5*, 1515–1539.

Roy, B., and Jacobson, A. (2013). The intimate relationships of mRNA decay and translation. Trends Genet. *29*, 691–699.

Schmid, M., and Jensen, T.H. (2018). Controlling nuclear RNA levels. Nat. Rev. Genet. *19*, 518–529.

Schmidt, C., Kowalinski, E., Shanmuganathan, V., Defenouillère, Q., Braunger, K., Heuer, A., Pech, M., Namane, A., Berninghausen, O., Fromont-Racine, M., et al. (2016). The cryo-EM structure of a ribosome-Ski2-Ski3-Ski8 helicase complex. Science *354*, 1431–1433.

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics *27*, 863–864.

Simms, C.L., Yan, L.L., and Zaher, H.S. (2017). Ribosome collision is critical for quality control during no-go decay. Mol. Cell *68*, 361–373.

Slevin, M.K., Meaux, S., Welch, J.D., Bigler, R., Miliani de Marval, P.L., Su, W., Rhoads, R.E., Prins, J.F., and Marzluff, W.F. (2014). Deep sequencing shows multiple oligouridylations are required for 3′ to 5′ degradation of histone mRNAs on polyribosomes. Mol. Cell *53*, 1020–1030.

Sohrabi-Jahromi, S., Hofmann, K.B., Boltendahl, A., Roth, C., Gressel, S., Baejen, C., Soeding, J., and Cramer, P. (2019). Transcriptome maps of general eukaryotic RNA degradation factors. eLife *8*, e47040.

Team, R.C. (2013). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Tesina, P., Heckel, E., Cheng, J., Fromont-Racine, M., Buschauer, R., Kater, L., Beatrix, B., Berninghausen, O., Jacquier, A., Becker, T., and Beckmann, R. (2019). Structure of the 80S ribosome-Xrn1 nuclease complex. Nat. Struct. Mol. Biol. *26*, 275–280.

Thandapani, P., Song, J., Gandin, V., Cai, Y., Rouleau, S.G., Garant, J.-M., Boisvert, F.-M., Yu, Z., Perreault, J.-P., Topisirovic, I., and Richard, S. (2015). Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. eLife *4*, e06234.

Travis, A.J., Moody, J., Helwak, A., Tollervey, D., and Kudla, G. (2014). Hyb: a bioinformatics pipeline for the analysis of CLASH (crosslinking, ligation and sequencing of hybrids) data. Methods *65*, 263–273.

Tuck, A.C., and Tollervey, D. (2013). A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. Cell *154*, 996–1009.

Tuck, A.C., Natarajan, K.N., Rice, G.M., Borawski, J., Mohn, F., Rankova, A., Flemr, M., Wenger, A., Nutiu, R., Teichmann, S., and Bühler, M. (2018). Distinctive features of lincRNA gene expression suggest widespread RNA-independent functions. Life Sci. Alliance *1*, e201800124.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

van Hoof, A., Staples, R.R., Baker, R.E., and Parker, R. (2000). Function of the ski4p (Csl4p) and Ski7p proteins in 3′-to-5′ degradation of mRNA. Mol. Cell. Biol. *20*, 8230–8243.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc. Natl. Acad. Sci. USA *101*, 11269–11274.

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, R8.

Welte, T., Tuck, A.C., Papasaikas, P., Carl, S.H., Flemr, M., Knuckles, P., Rankova, A., Bühler, M., and Großhans, H. (2019). The RNA hairpin binder TRIM71 modulates alternative splicing by repressing MBNL1. Genes Dev. *33*, 1221–1235.

Weren, R.D., Venkatachalam, R., Cazier, J.B., Farin, H.F., Kets, C.M., de Voer, R.M., Vreede, L., Verwiel, E.T., van Asseldonk, M., Kamping, E.J., et al. (2015). Germline deletions in the tumour suppressor gene FOCAD are associated with polyposis and colorectal cancer development. J. Pathol. *236*, 155–164.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag).

Widner, W.R., and Wickner, R.B. (1993). Evidence that the SKI antiviral system of Saccharomyces cerevisiae acts by blocking expression of viral mRNA. Mol. Cell. Biol. *13*, 4331–4341.

Zhang, E., Khanna, V., Dacheux, E., Namane, A., Doyen, A., Gomard, M., Turcotte, B., Jacquier, A., and Fromont-Racine, M. (2019). A specialised SKI complex assists the cytoplasmic RNA exosome in the absence of direct association with ribosomes. EMBO J. *38*, e100640.

Zhao, C., Datta, S., Mandal, P., Xu, S., and Hamilton, T. (2010). Stress-sensitive regulation of IFRD1 mRNA decay is mediated by an upstream open reading frame. J. Biol. Chem. *285*, 8552–8562.

Zinder, J.C., and Lima, C.D. (2017). Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors. Genes Dev. *31*, 88–100.

Zou, S., Chang, J., LaFever, L., Tang, W., Johnson, E.L., Hu, J., Wilk, R., Krause, H.M., Drummond-Barbosa, D., and Irusta, P.M. (2011). Identification of dAven, a Drosophila melanogaster ortholog of the cell cycle regulator Aven. Cell Cycle *10*, 989–998.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Mouse anti-FLAG M2 | Sigma | Cat#F1804 |
| Anti-FLAG M2 Dynabeads | Sigma | Cat#M8823 |
| Dynabeads M280 streptavidin-coated beads | Thermo Fisher | Cat#11206D |
| Dynabeads Protein G | Thermo Fisher | Cat# 10004D |
| Streptavidin-HRP | Sigma | Cat#S2438 |
| Rabbit anti-AVEN | ProScience | Cat#2417 |
| Rabbit anti-MTR4 | ThermoFisher Scientific | Cat#PA557927 |
| Rabbit anti-ATF4 D4B8 | Cell Signaling | Cat#mAb11815 |
| Rabbit anti-Phospho-eIF2α Ser51 D9G8 | Cell Signaling | Cat#mAb3398 |
| Rabbit anti-eIF4E | Bethyl Laboratories | Cat#A301-154A |
| Rat anti-tubulin clone YL1/2 | Abcam | Cat#ab6160 |
| Rat anti-HA | Roche | Cat#11867423001 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| T4 DNA Ligase | Sigma | Cat#10716359001 |
| DMEM | GIBCO | Cat#21969-035 |
| Non-essential amino acids | GIBCO | Cat#11140035 |
| 100 mM Sodium pyruvate | GIBCO | Cat#11360070 |
| 200 mM L-glutamine | GIBCO | Cat#25030024 |
| Fetal bovine serum | GIBCO | Cat#10270106 |
| Beta-mercaptoethanol | Sigma | Cat#M-7522 |
| Gelatin | Sigma | Cat#G-1890 |
| Trypsin-EDTA | GIBCO | Cat#25300-054 |
| Dulbecco's PBS | GIBCO | Cat#14190 |
| Trypsin (TPCK-treated) | Sigma | Cat#T8802 |
| OptiMEM | GIBCO | Cat#31985070 |
| Lipofectamine 3000 Transfection kit | Invitrogen | Cat#L3000015 |
| cOmplete Protease Inhibitor Cocktail | Roche | Cat#11836145001 |
| Proteinase K | Roche | Cat#3115879001 |
| SuperScript III | Life Technologies | Cat#18080085 |
| 3xFLAG peptide | Sigma | Cat#F4799-25MG |
| RNace-It Ribonuclease Cocktail | Agilent | Cat#400720 |
| TSAP Thermosensitive Alkaline Phosphatase | Promega | Cat#M9910 |
| RNasin Ribonuclease Inhibitor | Promega | Cat#N2115 |
| Recombinant RNasin Ribonuclease Inhibitor | Promega | Cat#N2511 |
| miR-cat 33 conversion oligo pack | IDT | N/A |
| T4 RNA Ligase 1 (ssRNA Ligase) | NEB | Cat#M0204L |
| T4 PNK, T4 polynucleotide kinase | NEB | Cat#M0201L |
| Hybond-C Extra membrane | GE Healthcare | Cat#RPN303E |
| Kodak BioMax MS autoradiography film | Kodak | Cat#8222648 |
| MetaPhor agarose | Lonza | Cat#50180 |
| NuPAGE 4–12% (wt/vol) polyacrylamide Bis-Tris gels | Life Technologies | Cat#NP0335 |
| NuPAGE LDS sample buffer 4 × | Life Technologies | Cat#NP0007 |
| NuPAGE SDS-MOPS running buffer | Life Technologies | Cat#NP0001 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| NuPage transfer buffer | Life Technologies | Cat#NP00061 |
| MinElute Gel extraction kit | QIAGEN | Cat#28604 |
| Proteinase K | Roche | Cat#03115836001 |
| RNase H | NEB | Cat#M0297L |
| TaKaRa long and accurate (LA) Taq | Clontech | Cat#RR002M |
| γ32P-ATP 0.5 mCi 18.5 MBq Spec act. > 6000 Ci/mmol | Hartman | Cat#SRP-501 |
| NEBNext® High-Fidelity 2X PCR Master Mix | NEB | Cat#M0541 |
| 4-hydroxytamoxifen | Sigma | Cat#H6278 |
| Puromycin | Sigma | Cat#P8833 |
| Cycloheximide | Sigma | Cat#C7698 |
| Harringtonine | LKT Laboratories | Cat#H0169 |
| Immobilon Western Chemiluminiscent HRP Substrate | Merck Millipore | Cat#WBKLS0500 |
| Thapsigargin | Invitrogen | Cat#T7459 |
| **Critical Commercial Assays** | | |
| ScriptSeq RNA-Seq Library Prep Kit | NEB | Cat#E7645 |
| Agilent Absolutely RNA Miniprep Kit | Epicenter | Cat#SSV21106 |
| TruSeq RNA Library Prep Kit v2 | Illumina | Cat#RS-122-2001 |
| miRNeasy RNA Extraction kit | QIAGEN | Cat#217004 |
| Ribo-Zero Gold rRNA Removal Kit | Illumina | Cat#MRZG12324 |
| PrimeScript RT-PCR Kit | Takara Bio | Cat#RR036A-1 |
| Qubit dsDNA HS Assay Kit | Thermo Fisher | Cat#Q32854 |
| SsoAdvanced SYBR Green Supermix | Bio-Rad | Cat#172-5274 |
| **Deposited Data** | | |
| CRAC | This paper | GEO: GSE134020 |
| Ribosome profiling (monosome and disome profiling) | This paper | GEO: GSE134020 |
| RNA-seq | This paper | GEO: GSE134020 |
| Human ribosome structure | PDB | PDB #4UG0 |
| Mouse pre-rRNA sequence | Grozdanov et al., 2003 | GenBank BK000964 |
| *Mus musculus* GRCm38/mm10 genome assembly, Mus_musculus.GRCm38.75.dna.primary_assembly.fa | Ensembl | ftp://ftp.ensembl.org/pub/release-75/fasta/mus_musculus/dna/ |
| Gene annotations: Gencode M16 = Ensembl 91 (GRCm38) (including tRNAs and Appris isoforms) | Gencode | https://www.gencodegenes.org/mouse/release_M16.html |
| Original uncropped western blot images | This study | N/A |
| **Experimental Models: Cell Lines** | | |
| *Rosa26*$^{Cre\text{-}ERT2/-}$ | Flemr and Bühler, 2015 | cMB052 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ | Ostapcuk et al., 2018 | cMB063 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Xrn1*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | This study | cMB315 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Aven*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | This study | cMB323 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | This study | cMB331 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Mtr4*$^{1xFlag\text{-}Avi/1xFlag\text{-}Avi}$ | Tuck et al., 2018 | cMB376 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Rps10*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ | This study | cMB395 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$*Focad*$^{-/-}$ | This study | cMB396 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$*Focad*$^{-/-}$ | This study | cMB397 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$*Aven*$^{-/-}$ | This study | cMB399 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$*Aven*$^{-/-}$ | This study | cMB400 |
| *Rosa26*$^{Cre\text{-}ERT2/-}$ *Skiv2l*$^{fl/fl}$ | This study | cMB434 |
| *Rosa26*$^{Cre\text{-}ERT2/-}$ *Skiv2l*$^{fl/fl}$ | This study | cMB435 |

(*Continued on next page*)

**Continued**

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Aven*$^{-/-}$ *Skiv2l*$^{fl/fl}$ | This study | cMB471 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Aven*$^{-/-}$ *Skiv2l*$^{fl/fl}$ | This study | cMB472 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Mtr4*$^{3xFlag\text{-}Avi/3xFlag\text{-}Avi}$ | Tuck et al., 2018 | cMB503 |
| *Rosa26*$^{Cre\text{-}ERT2/BirA\text{-}V5}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Dis3l*$^{2xHA\text{-}FKBP12(F36V)/\ 2xHA\text{-}FKBP12(F36V)}$ | This study | cMB510 |
| Oligonucleotides | | |
| qPCR primers, see Table S1 | This paper | N/A |
| Donor oligonucleotides for genome editing, see Table S1 | This paper | N/A |
| 5′ adapters for CRAC (barcodes highlighted): | | N/A |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrUrArArGrC | L5Aa | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrArUrUrArGrC | L5Ab | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrGrCrGrCrArGrC | L5Ac | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrCrGrCrUrUrArGrC | L5Ad | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrArGrArGrC | L5Ba | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrGrUrGrArGrC | L5Bb | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrCrArCrUrArGrC | L5Bc | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrUrCrUrCrUrArGrC | L5Bd | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrCrUrArGrC | L5Ca | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrUrGrGrArGrC | L5Cb | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrArCrUrCrArGrC | L5Cc | IDT custom synthesis |
| /5InvddT/ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrUrNrNrNrNrGrArCrUrUrArGrC | L5Cd | IDT custom synthesis |
| AATGATACGGCGACCACCGAGATCTACACT CTTTCCCTACACGACGCTCTTCCGATCT | P5 | IDT custom synthesis |
| CAAGCAGAAGACGGCATACGAGATCGGTCT CGGCATTCCTGGCCTTGGCACCCGAGAATTCC | PE | IDT custom synthesis |
| Software and Algorithms | | |
| STAR 2.5.0a | Dobin et al., 2013 | N/A |
| Bedtools 2.26.0 | Quinlan, 2014 | N/A |
| Samtools 1.6 | Li et al., 2009 | N/A |
| R version 3.5.1 Patched (2018-11-02 r75543) | R Core Team, 2013 | https://www.r-project.org/ |
| ggplot2 3.1.0 | Wickham, 2016 | N/A |
| FASTX Toolkit 0.0.14 | | http://hannonlab.cshl.edu/fastx_toolkit/ |
| pyCRAC | Webb et al., 2014 | N/A |
| prinseq-lite-0.20.4 | Schmieder and Edwards, 2011 | N/A |
| bowtie2-2.3.4.1 | Langmead and Salzberg, 2012 | N/A |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| DESeq2 | Love et al., 2014 | N/A |
| RNAfold 2.1.5 | Lorenz et al., 2011 | N/A |
| cutadapt | Martin, 2011 | N/A |
| StringTie 1.3.3b | Pertea et al., 2015 | N/A |
| edgeR v3.16.5 | Robinson et al., 2010 | N/A |

## LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Marc Bühler (marc.buehler@fmi.ch). All unique reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

Male 129 × C57BL/6 mouse embryonic stem cells (mESC) (Mohn et al., 2008) were grown in serum/LIF media (DMEM (GIBCO 21969-035) supplemented with 15% fetal bovine serum (GIBCO 10270106), 2 mM L-glutamine (GIBCO 25030024), 1x non-essential amino acids (GIBCO 11140035), 1 mM sodium pyruvate (GIBCO 11360070), 0.1 mM 2-mercaptoethanol (Sigma M-7522), 50 mg $ml^{-1}$ penicillin, 80 mg $ml^{-1}$ streptomycin and homemade LIF) at 37 °C in 5% CO2. Cells were cultured on dishes coated with 0.1% gelatin (Sigma G1890).

## METHOD DETAILS

### Generation of endogenously tagged cell lines

Endogenous gene tagging with a 3xFLAG-AviTag was performed in mES 129 × C57BL/6 cells expressing BirA ligase and CreERT2 from the *Rosa26* locus (cMB063) (Ostapcuk et al., 2018), using TALEN or CRISPR-Cas9 homology-directed repair with single-stranded oligodeoxynucleotide (ssODN) donor templates encoding the tag, flanked by 5′ and 3′ homology arms. The ssODNs donors were synthetized as ultramers by Integrated DNA Technologies. N-terminally tagged *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 8F (cMB331) and *Aven*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 2B (cMB323) were generated using TALENs and Cas9/gRNA, respectively, cutting near the start codon. *Xrn1*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 4F (cMB315) was C-terminally tagged using Cas9/gRNA cutting near the stop codon. N-terminally tagged *Mtr4* cell lines (cMB376 and cMB503) were previously described (Tuck et al., 2018). C-terminally tagged *Rps10*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ clone 4E (cMB395) was generated using Cas9/gRNA cutting near the stop codon. N-terminally tagged *Dis3l*$^{2xHA\text{-}FKBP12(F36V)/\ 2xHA\text{-}FKBP12(F36V)}$ (cMB510) was generated in the *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ (cMB331) background using Cas9/gRNA cutting near the start codon. For homology-directed repair, the donor sequence encoding the 2xHA-FKBP12(F36V) tag, flanked by ~550bp *Dis3l* 5′ and 3′ homology arms was cloned into a pBLU plasmid and transfected together with the Cas9/gRNA. All clones were screened for homozygous integration of the tag by PCR and Sanger sequencing and expression of the fusion proteins was confirmed by western blot with an anti-FLAG or anti-HA antibody. Biotinylation of the tag was verified by western blot using streptavidin-HRP. A full list of genome-edited cell lines together with TALENs, gRNAs and donor ssODN ultramer sequences can be found in Table S1.

### Generation of straight KO cell lines

*Aven*$^{-/-}$ clones 4H (cMB399) and 6G (cMB400) were generated in a *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ background (cMB331) using Cas9/gRNAs targeting *Aven* exon 3 and exon 6 (last exon), resulting in a deletion of approximately 5.7 kb. *Focad*$^{-/-}$ clones 2F (cMB396) and 4B (cMB397) were generated in a *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ background (cMB331) with Cas9/gRNAs targeting intron 2 and intron 4. The resulting deletion of approximately 7.3 kb introduces a frameshift in exon 5. Homozygous knockout clones were screened by PCR and Sanger sequencing and deletion was confirmed by western blot or RT-qPCR. See also Table S1.

### Generation of conditional KO cell lines

*Skiv2l*$^{fl/fl}$ cell lines were generated in a 129 × C57BL/6 WT background expressing a CreERT2 recombinase fusion from the *Rosa26* locus (cMB052) as well as in *Aven*$^{-/-}$ cells where *Skiv2l* is endogenously tagged (cMB399). A plasmid expressing Cas9 and gRNAs targeting *Skiv2l* intron 10 and intron 17 was co-transfected with ssODN containing homology arms and LoxP sites for integration. Recombination of the LoxP sites eliminates exons 11-17 containing the catalytic DExH box and results in out-of-frame translation of the last 18 exons. Clones with homozygous insertions of LoxP sites in both intron 10 and intron 17 were screened by PCR and Sanger sequencing. Proper recombination of the LoxP sites was tested by RT-qPCR, or western blot, following treatment with 0.1 μM 4-hydroxytamoxifen (4-OHT) (Sigma) for 2, 4 or 6 days. See also Table S1.

## Transfections

For genome editing with CRISPR-Cas9, gRNAs were cloned into the SpCas9-2A-mCherry vector (Knuckles et al., 2017). To generate endogenously tagged *Xrn1* (cMB315), *Aven* (cMB323) and *Rps10* (cMB395), cells were transfected with 1000 ng SpCas9-2A-mCherry, 1400 ng ssODN donor and 100 ng pRRE GFP homologous recombination reporter (Flemr and Bühler, 2015). mCherry and GFP double-positive cells were FACS-sorted 24 hours after the transfection and seeded sparsely (10,000 cells) on 10 cm plates for clonal expansion. After 5-7 days, colonies were individually picked into 96-well plates, expanded and genotyped by PCR. Cells with proper in-frame homozygous insertions of the tag were further confirmed by Sanger sequencing and western blot.

For endogenous tagging of *Skiv2l* (cMB331) with TALENs, cells were transfected with 400 ng of each TALEN, 1000 ng ssODN donor and 100 ng of pRRP puromycin recombination reporter (Flemr and Bühler, 2015). 24 hours post-transfection, the cells were selected with 2 μg/ml puromycin for 28 hours and surviving cells were plated at clonal densities as described above. *Skiv2l*$^{fl/fl}$ cell lines were generated by transfecting 450 ng of each SpCas9-2A-mCherry gRNA plasmid, 500 ng of each LoxP ssODN donor and 50 ng of each pRRP puromycin reporter and selection with 2 μg/ml puromycin.

To create *Aven*$^{-/-}$ (cMB399 and cMB400) and *Focad*$^{-/-}$ (cMB396 and cMB397), cells were transfected with 500 ng of each SpCas9-2A-mCherry gRNA vector and 50 ng of each pRRP or pRRE-GFP reporter. *Aven*$^{-/-}$ cells were selected on 2 μg/ml puromycin and *Focad*$^{-/-}$ cells were selected by FACS-sorting mCherry-GFP double-positives.

To generate endogenously tagged *Dis3l*$^{2xHA\text{-}FKBP12(F36V)/\ 2xHA\text{-}FKBP12(F36V)}$ (cMB510), *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ (cMB331) cells were transfected with 500 ng SpCas9-2A-mCherry, 700 ng pBLU 2xHA-FKBP12(F36V) donor plasmid and 100 ng pRRP puromycin reporter. The cells were selected with 2 μg/ml puromycin and genotyped as described above. All transfections were carried out with Lipofectamine 3000 reagent at 3 μl per 1 μg of total DNA in OptiMem media. Approximately 500,000 cells were used for each transfection.

## RNA sequencing

Total RNA was extracted from ~80% confluent 6 cm dishes using the Agilent Absolutely RNA Miniprep Kit with on-column DNase digestion. After ribosomal RNA depletion with the Illumina Ribozero kit, libraries were constructed using either ScriptSeq v2 or TruSeq v2 kits and sequenced on an Illumina HiSeq2500 platform (50 nt single-end reads). Total RNA from *Skiv2l*$^{fl/fl}$ conditional knockouts was extracted after culturing the cells in media supplemented with 0.1 μM 4OHT for 0, 2, 4 or 6 days to induce *Skiv2l* knockout.

To measure transcriptome-wide RNA half-lives, 300,000 mESCs were seeded per well of a six-well dish and grown for 48 h in serum + LIF medium. The medium was replaced by fresh medium with 5 μM actinomycin D (from a 5 mg/ml stock in DMSO) and the cells were incubated for 120, 240 or 360 min. A mock treatment (360 min) was included, using medium with the same amount of DMSO but no actinomycin D. After the indicated times, cells were washed twice with 37°C PBS and RNA extracted using the Agilent Absolutely RNA Miniprep kit. ERCC RNA spike-ins were added to the lysis buffer (1.7 μL of a 1:10 dilution per sample) before it was added to the cells. Three technical replicates were performed for each cell line, treatment and time point.

## CRAC

CRAC was performed as described in (Tuck et al., 2018), with minor modifications, and is described in full here:

mESCs were grown in 2x 15-cm dishes to ~80% confluency, dishes washed 2x with PBS, the PBS removed, then cells crosslinked on ice (with dishes facing up) in a Stratagene Stratalinker 2400 (400 mJ cm$^{-2}$). Cells were lysed by incubating with 5 mL of either TN150+NP40 (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 0.5% Nonidet P40 substitute and 1x cOmplete Protease Inhibitor Cocktail), or of RIPA (50 mM Tris-HCl pH 7.8, 300 mM NaCl, 1.0% Nonidet P40 substitute, 0.1% SDS, 10% (v/v) glycerol, 0.5% sodium deoxycholate, 1 mM beta-mercaptoethanol, 1x cOmplete Protease Inhibitor Cocktail), as indicated in Table S2. The harsher buffer (RIPA) was initially used to ensure complete extraction, but as this can reduce FLAG binding we later switched to a milder buffer (TN150), which did not affect library content. The cells were further disrupted using a cell scraper then lysates collected and centrifuged (6500 xg for 20 min at 4°C). Supernatants were frozen in liquid nitrogen and stored at −80°C.

Lysates were thawed on ice and incubated with 100 μL anti-FLAG M2 magnetic beads overnight. The supernatant was discarded and beads washed 3x with 1 mL TN150 (50 mM Tris-HCl pH 7.8, 150 mM NaCl, 0.1% Nonidet P40 substitute). Protein:RNA complexes were eluted by incubating beads in 1.5 mL TN150 supplemented with 5 mM beta-mercaptoethanol and 0.3 mg/ml 3xFLAG peptide, rotating at 4°C for 2 hr. The eluate was then incubated with 50 μL Dynabeads M-280 Streptavidin, rotating at 4°C overnight. Beads were washed 2x in TN600 (50 mM Tris-HCl pH 7.8, 600 mM NaCl, 0.1% Nonidet P40 substitute, 5 mM beta-mercaptoethanol) and 2x in TN150 supplemented with 5 mM beta-mercaptoethanol. RNA was fragmented by incubating beads in 500 μL TN150 supplemented with 5 mM beta-mercaptoethanol and 1 μL of 0.1 U diluted RNace-IT. After 4 min at 37°C, the RNases were denatured by replacing the solution with 400 μL WBI (50 mM Tris-HCl pH 7.8, 300 mM NaCl, 0.1% Nonidet P40 substitute, 5 mM beta-mercaptoethanol and 4.0 M guanidine hydrochloride). The beads were washed 2x in WBI, then 3x in 400 μL 1xPNK (50 mM Tris-HCl pH 7.8, 10 mM $MgCl_2$, 0.5% Nonidet P40 substitute, 5 mM beta-mercaptoethanol).

The following four enzymatic reactions were then performed in 80 μL 1xPNK buffer (omitting the Nonidet P40 substitute), to ligate 3′ and 5′ adapters onto RNA fragments. After each reaction, beads were washed 1x in WBI and 3x in 1xPNK:

(i) Alkaline phosphatase treatment (30 min, 37°C): 8 U TSAP, 80 U RNasIN.

(ii) 3′ linker ligation (overnight, 16°C): 0.1 nmol miRCat-33 DNA linker, 40 U T4 RNA Ligase 1, 80 U RNasIN, 12.5% (v/v) PEG8000.
(iii) 5′ phosphorylation (1 hr, 37°C): 40 U T4 PNK, 2 μL γ32P-ATP (after 30 min, add 1 μL 100 mM rATP and an additional 20 U T4 PNK).
(iv) 5′ linker ligation (overnight, 16°C): 0.2 nmol 5′ linker, 40 U T4 RNA Ligase 1, 1.25 mM rATP, 80 U RNasIN, 12.5% (v/v) PEG8000.

After the final reaction, beads were washed 3x in WBII (50 mM Tris-HCl pH 7.8, 50 mM NaCl, 0.1% Nonidet P40 substitute, 5 mM beta-mercaptoethanol), resuspended in 30 μL 1x NuPAGE LDS sample buffer, heated at 95°C for 2 min, and the eluate quickly removed and loaded onto a NuPAGE 4%–12% polyacrylamide gel. The gel was run at 100 V for ~1 hr, then protein:RNA complexes transferred to Hybond-C extra nitrocellulose membrane (Amersham) at 150 V for 1.5 hr using a wet transfer system and NuPAGE transfer buffer with 15% methanol. The membrane was then briefly dried, exposed to BioMax MS film (4 hr to overnight) and the region corresponding to the protein:RNA complex cut out.

The membrane slice was then incubated in 400 μL WBII with 1% (w/v) SDS, 5 mM EDTA and 100 μg Proteinase K at 55°C for 2 hr. The solution was then removed to another tube, 50 μl 3M NaAc pH 5.2 and 500 μl of 1:1 phenol:chloroform mix added, and the mixture vortexed then centrifuged at 14,000 xg for 20 min. The top phase was transferred into a new tube and 1 mL ethanol and 20 μg glycogen added. The solution was stored at −20°C overnight to precipitate RNA, then centrifuged at 14,000 xg for 1 hr. The pellet was washed once with 70% ethanol and allowed to briefly air dry, before resuspending in 11 μL water + 1 μL 10 μM miRCat-33 RT oligo + 1 μL 10 mM dNTP mix. The solution was heated to 80°C for 3 min, snap cooled on ice for 5 min, then the following mix added: 4 μL 5x first strand buffer (SSIII kit) + 1 μL 100 mM DTT (SSIII kit) + 1 μL recombinant RNasIN. After incubating for 3 min at 50°C, 200 U of SuperScript III was added and the reverse transcription allowed to proceed for 1 hr at 50°C. The reaction was stopped by heating to 65°C for 15 min, then RNA digested with 10 U RNase H at 37°C for 30 min. PCR reactions (80 μL) were then prepared, each with 2 μL cDNA, 10 pmol P5, 10 pmol PE, 12.5 nmol each dNTP and 2.5 U LA Taq. Typically, we ran five PCR reactions and then concentrated the products by ethanol precipitation before resolving on a 3% metaphor agarose gel in 0.5x TBE. A smear corresponding to the size of the two adapters plus inserts (total size ~120-300 bp) was then excised, and DNA extracted using the MinElute gel extraction kit, eluting in 20 μL water. If the experiment was successful, we repeated the PCRs with the remaining half of the cDNA.

The above CRAC protocol is referred to as the "long" protocol. For some samples (indicated in Table S2), a shorter version of the protocol was used, which did not affect library content. The shorter version omits radiolabelling (using cold rATP instead), SDS-PAGE and transfer to nitrocellulose. Instead, after 3′ linker ligation, beads were washed and added directly to 400 μL WBII with 1% (w/v) SDS, 5 mM EDTA and 100 μg Proteinase K. This version of the CRAC protocol is referred to as the "short" protocol (indicated in Table S2).

### Translation inhibition experiments for CRAC

Cells grown to ~80% confluency on 15cm dishes were incubated with media supplemented with either 100 μg/mL cycloheximide or 5 μM harringtonine for 30 min at 37°C. Cells were then washed twice with PBS containing the same concentration of the corresponding inhibitors. PBS was removed after the last wash and the cells were cross-linked on ice, with the dishes facing up, in a Stratagene Stratalinker 2400 (400 mJ·cm−2) and processed for CRAC as described above.

### Ribosome profiling

Cells were harvested (without cycloheximide pretreatment) and flash-frozen in liquid nitrogen. From the cell pellets, lysates were prepared and ribosome-protected mRNA fragments were generated by RNase I digestion as previously described (Janich et al., 2015). For the excision of footprints from 15% urea-polyacrylamide gels, single strand RNA oligonucleotides of 26 nt and 34 nt (for monosome footprints) and of 52 nt and 69 nt (for disome footprints) served as size markers for excision of footprints. After fragment purification with miRNeasy RNA Extraction kit, 5μg fragmented RNA was used for ribosomal RNA removal using Ribo-Zero Gold rRNA Removal Kit according to Illumina's protocol for TruSeq Ribo Profile (RPHMR12126 Illumina).

Sequencing libraries were generated according to Illumina's TruSeq Ribo-Profile protocol with minor modifications. Monosomes and disomes were treated as independent libraries. cDNA fragments were separated on a 10% urea-polyacrylamide gel and gel slices between 70-80 nt for monosomes and 97-114 nt for disomes were excised. The PCR-amplified libraries were size selected on an 8% native polyacrylamide gel. Monosome libraries were at ~150 bp and disome libraries at ~180 bp.

Parallel RNA-seq libraries were prepared essentially following the Illumina protocol (Janich et al., 2015); briefly, after total RNA extraction using miRNeasy RNA Extraction kit, ribosomal RNA was depleted using Ribo-Zero Gold rRNA Kit, and sequencing libraries generated from the heat-fragmented RNA as described (Janich et al., 2015). All libraries were sequenced on Illumina HiSeq 2500.

### Western Blotting

Cells were lysed for 30 min on ice in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Triton-X, 0.5 mM EDTA, 5% glycerol, 1x protease inhibitor cocktail (Roche) and 1 mM DTT. Lysates were clarified by centrifugation at 16,000 xg for 10 min at 4°C and protein concentration was measured using the BioRad protein assay. Approximately 20 μg of total protein extract was resolved on

NuPAGE-Novex Bis-Tris 4%–12% gradient gels (Thermo Fisher NP0322BOX), transferred semi-dry to a polyvinylidene fluoride (PVDF) membrane, blocked in 5% non-fat milk in TBS+0.05% Tween (TBST) for 30 min at room temperature and incubated with primary antibodies at 4 °C overnight. The following primary antibodies were used for western blotting: mouse anti-Flag (1:1,000, Sigma clone M2), rabbit anti-AVEN (1:2,000, ProScience 2417), rabbit anti-ATF4 (1:1,000, Cell Signaling D4B8 mAb11815), rabbit anti- Phospho-eIF2α Ser51 (1:1,000, Cell Signaling D9G8 mAb3398), rabbit anti-eIF4E (1:1,000, Bethyl A301-154A) and rat anti-tubulin (1:5,000, Abcam clone YL1/2). Following incubation with corresponding HRP-conjugated secondary antibodies, signal was visualized using Immobilon Western Chemiluminiscent HRP Substrate. To detect biotinylated proteins, after transfer, membranes were blocked in 2% bovine serum albumin (BSA) in TBST for 30 min and incubated with HRP-conjugated streptavidin (Strep-HRP) diluted 1:10,000 in 2% BSA-TBST for 30 min at room temperature. For detection of ATF4 and Phospho-eIF2α Ser51, membranes were first probed for ATF4, stripped in 25 mM Glycine, pH 2 and 1% SDS for 5 min at room temperature, rinsed with TBST, blocked in 5% non-fat milk TBST for 30 min and re-probed for Phospho-eIF2α Ser51.

### Co-immunoprecipitations

*Dis3l*$^{2xHA\text{-}FKBP12(F36V)/2xHA\text{-}FKBP12(F36V)}$ (cMB510) cells grown to ~80% confluency in a 10 cm dish were trypsinized, collected in media and washed twice with PBS. The cells were lysed for 40 min at 4°C in 500 μl lysis buffer (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2.5 mM $MgCl_2$, 0.5% NP-40), supplemented with 1X protease inhibitor cocktail (Roche). Lysates were clarified by centrifugation at 16,000 g for 5 min and mixed with 30 μl Protein-G Dynabeads (Thermo Fisher 10004D) coupled to 2 μg anti-HA antibody (Roche 11867423001). The sample was incubated for 1 hour at 4°C on a rotating wheel. The beads were then washed four times in wash buffer (10mM Tris-HCl, pH 7.4, 150 mM NaCl, 2.5 mM $MgCl_2$, 0.1% NP-40), resuspended in 60 μl 1X Bolt LDS Sample Buffer (Thermo Fisher B0007) and incubated at 85°C for 5 min to elute captured proteins from the beads. Following this, 2% of the input and 30% of the IP material were resolved on a NuPAGE-Novex Bis-Tris 4%–12% gradient gel (Thermo Fisher NP0322BOX), transferred semi-dry to a polyvinylidene fluoride (PVDF) membrane, blocked in 5% non-fat milk in TBS+0.05% Tween (TBST) for 30 min at room temperature and incubated with primary antibodies at 4 °C overnight. The following primary antibodies were used for western blotting: rat anti-HA (1:1,000, Roche 11867423001), mouse anti-Flag (1:1,000, Sigma clone M2), rabbit anti-Mtr4 (1:1,000, Thermo Fisher PA5-57927).

### Affinity purification for LC–MS/MS

For tandem FLAG-streptavidin affinity purification, two confluent 15 cm dishes seeded with equal number *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ cells or the corresponding untagged parental line were harvested by trypsinization, washed twice in PBS and lysed 2 hours to overnight at 4°C in whole cell lysis buffer (10 mM Tris-HCl pH 7.4, 150 mM KCl, 2.5 mM $MgCl_2$, 0.5% NP-40), supplemented with 1X protease inhibitor cocktail, 50 units benzonase and 10 μg RNase A. Lysates were clarified by centrifugation at 16,000 g for 15 min and incubated with 20 μl anti-FLAG M2 Dynabeads for 4 hours at 4°C. After washing the FLAG beads three times with wash buffer (10mM Tris-HCl (pH 7.4), 150 mM KCl. 2.5 mM $MgCl_2$, 0.1% NP-40), proteins were eluted from the beads three times for 15 min at 4°C with 100 μg/ml 3xFLAG peptide diluted in 150 μl wash buffer. The eluates were combined and incubated with 20 μl M-280 Streptavidin Dynabeads for 2 hours at 4°C, washed four times in wash buffer and two times in wash buffer without NP-40. For mass spectrometry analysis, captured proteins were digested with trypsin directly on the streptavidin beads. High-salt tandem FLAG-strep affinity purifications were essentially carried out as described above with the following modifications: cells from two confluent 10 cm dishes were lysed in buffer containing 350 mM KCl. For single-step streptavidin pull-downs, the FLAG purification step was omitted and total cell lysates from two confluent 10 cm dishes were applied directly to streptavidin beads. Every affinity purification experiment contained three separate technical replicates for each cell line.

### Mass spectrometry analysis

Peptides generated by trypsin digestion (see 'Affinity purification for LC–MS/MS') were acidified with 0.8% TFA (final) and analyzed by LC–MS/MS on an EASY-nLC 1000 with a two column set-up (Thermo Scientific). The peptides were applied onto a peptide trap (Acclaim PepMap 100, 75 μm × 2 cm, C18, 3 μm, 100 Å) in 0.1% formic acid, 2% acetonitrile in $H_2O$ at a constant pressure of 80 MPa. Using a flow rate of 150 nl $min^{-1}$, peptides were separated with a linear gradient of 2%–6% buffer B in buffer A in 3 min followed by an linear increase from 6 to 22% in 40 min, 22%–28% in 9 min, 28%–36% in 8 min, 36%–80% in 1 min and the column was finally washed for 14 min at 80% buffer B in buffer A (buffer A: 0.1% formic acid; buffer B: 0.1% formic acid in acetonitrile) on a 50 μm × 15 cm ES801 C18, 2 μm, 100 Å column (Thermo Scientific) mounted on a DPV ion source (New Objective) connected to a Orbitrap Fusion (Thermo Scientific). Data acquisition was performed using 120,000 resolution for the peptide measurements in the Orbitrap and a top T (3 s) method with HCD fragmentation for each precursor and fragment measurement in the ion trap following the manufacturer guidelines (Thermo Scientific).

Peptide identification was performed with MaxQuant version 1.5.3.8 using Andromeda as search engine (Cox et al., 2011). The mouse subset of the UniProt version 2015_01 combined with the contaminant DB from MaxQuant was searched and the protein and peptide FDR values were set to 0.05. All MaxQuant parameters can be found in Table S5.

Statistical analysis was done in Perseus (version 1.5.2.6) (Tyanova et al., 2016). Results were filtered to remove reverse hits, contaminants and peptides found in only one sample. Missing values were imputed and potential interactors were determined using t test and visualized by a volcano plot. Significance lines corresponding to an FDR of 0.05 and S0 (curve bend) between 0.2 and

2.0 are shown in the corresponding Figures. Results were exported from Perseus and visualized using statistical computing language R.

### Cell cycle analysis

Cells were synchronized at the G1/S boundary using a double-thymidine block. Briefly, 300,000 cells of each indicated cell line were seeded on 6-well plates and grown overnight in normal serum/LIF media, or media containing 0.1 μM 4-OHT to induce *Skiv2l* knockout where necessary. On the following day, the cells were switched to media supplemented with 2 mM Thymidine and cultured for 18 hours, released into the cell cycle for 9 hours by removal of the drug with three PBS washes and cultured in 2 mM thymidine media for another 18 hours. The cells were then released from the second block by three PBS washes and harvested at 0, 4 and 8 hours after thymidine withdrawal. For each sample, equal number of cells were fixed in ice-cold 70% ethanol and incubated overnight at 4°C. The cells were then permeabilized with PBS + 0.1% triton X200 for 2 min, stained with 1 μg/mL DAPI in PBS+0.1% triton X200 and DNA content was analyzed by flow cytometry. Histogram plots were generated using the FlowJo software.

### ATF4 induction with Thapsigargin

Approximately 120,000 cells per sample were seeded in 24-well plates and cultured in normal serum/LIF media overnight. On the following day the cells were switched to media supplemented with 200 mM Thapsigargin to induce the integrated stress response and upregulation of ATF4, and harvested for western blot analysis at 0, 0.5, 2 or 4 hours of incubation with the drug.

### RT-qPCR

Total RNA was extracted from mES cells with the Agilent Absolutely RNA Miniprep Kit and 500ng of RNA was reverse-transcribed using the PrimeScript RT-PCR Kit. qPCR was performed with SsoAdvanced SYBR Green Supermix on a CFX96 Real-Time PCR System (Bio-Rad) and relative RNA levels were calculated using the $\Delta C_t$ method and normalization to TBP mRNA abundance. A list of qPCR primers is provided in Table S1.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### CRAC data preprocessing and alignment

CRAC reads were preprocessed with the FASTX Toolkit 0.0.14. Adapters were removed with fastx_clipper, low quality bases trimmed/reads removed using fastq_quality_trimmer -t 25 and fastq_quality_filter -q 20 -p 90, and sequencing artifacts removed using fastx_artifacts_filter. Duplicate reads (including UMI) were collapsed, then pyCRAC (Webb et al., 2014) used to split samples by their inline barcodes and extract the UMIs . Low complexity regions were removed (''low complexity stripping'') from the 3′ end of sequences (defined as stretches of 2 nt more where 80% of positions are the same nucleotide, e.g., AAAAGAA), and prinseq-lite-0.20.4 (Schmieder and Edwards, 2011) used as an additional filter to remove low complexity reads (settings -lc_threshold 20 -lc_method dust). We then applied a set of filters to obtain uniquely mapping reads, and remove reads mapping to repeats or abundant non-coding RNAs (e.g., tRNA, snoRNA or rRNA). For this, reads were separately mapped with bowtie2-2.3.4.1 (Langmead and Salzberg, 2012) (settings–local -p 10 -a–very-sensitive) to three indexes:

Genome: mm10 genomic sequence.

Gencode non-coding RNAs: Gencode release M16 Mt_rRNA, Mt_tRNA, miRNA, rRNA, scRNA, snoRNA, sRNA, scaRNA and snRNA features from the file gencode.vM16.annotation.gtf, and predicted tRNAs from the file gencode.vM16.tRNAs.gtf, and a repeat-masked version of mouse pre-rRNA (Grozdanov et al., 2003).

Gencode mRNAs/lincRNAs: Gencode release M16 protein_coding and lincRNA features from the file gencode.vM16.annotation.gtf.

Reads were assigned as ''repeats/ncRNAs'' and excluded if they mapped best or equally well either to regions of the genome overlapping RepeatMasker repeats (downloaded from UCSC table browser, version 2012-02-07) or to Gencode non-coding RNAs. From the remaining reads, those mapping better or equally well to Gencode mRNA/lincRNAs as to the genome were extracted. A filter was then applied, by mapping these reads (bowtie2-2.3.4.1–local -p 10 -a–very-sensitive) to an index with one transcript isoform per protein-coding gene (the APPRIS principal isoform was selected, taking one at random if a gene possessed multiple; were refer to these as ''Appris transcripts''), and selecting reads with (i) a second best match score (if detected) < 0.8 times the best match score and (ii) a MAPQ $\geq$ 8. Reads were removed if the 5′ end was soft-clipped, and duplicate reads collapsed based on their UMIs and 5′ end mapping positions (retaining one read at random).

For comparing CRAC data with ribosome profiling, an identical procedure was used, except using a list of transcripts robustly detected in the ribosome profiling experiments instead of Appris transcripts. Where multiple transcripts were detected from the same gene, alignments were prioritized to the primary isoform if it could be determined (see ribosome profiling methods).

### CRAC quantification of non-templated 3′ tails

To identify reads containing non-templated 3′ tails, preprocessed reads identified as uniquely mapping to Appris transcripts were extracted, then filtered to retain those for which the 3′ adaptor could be identified and stripped. Bowtie2-2.3.4.1 was then used to align these reads (which were not subject to low complexity stripping from the 3′ end) to the genome, transcriptome and Appris transcripts (defined above), with the following settings, as described in (Travis et al., 2014): -D 20 -R 3 -N 0 -L 16–local -i S,1,0.50–score-min L,18,0–ma 1–np 0–mp 2,2–rdg 5,1–rfg 5,1. Tails were extracted by identifying examining alignments for 3′ soft clipping. Reads were required to align better to the transcriptome than to the genome (to prevent ''tails'' being identified that in fact correspond to exon-exon junctions). Only homopolymeric tails were analyzed, with no minimum length requirement.

### CRAC PCA, correlation matrix and tSNE

Filtered, uniquely mapping CRAC reads were counted for all mRNAs (using alignments to Appris transcripts, as defined above, and thus excluding reads mapping to introns). Each CRAC replicate was processed separately. For tSNE analysis, CRAC replicates were pooled, to give one dataset each for MTR4, SKIV2L and XRN1. These three count datasets were then normalized to the sum of the smallest dataset, and mRNAs retained with > 1 normalized count in all datasets, and > 10 normalized counts in at least one dataset. To obtain ''relative binding'' to MTR4, SKIV2L and XRN1, for each transcript the normalized counts were divided by its total normalized counts. Therefore, for each transcript, the ''relative binding'' of MTR4 + SKIV2L + XRN1 sums to 1. In parallel, to check that small differences in transcript levels between cell lines do not distort the analysis, relative binding values were normalized to rpkm values taken from RNA-seq analysis of the three cell lines.

### Differential SKIV2L binding analysis

As genes were differentially expressed in *Aven*$^{-/-}$ cells versus WT, an interaction model accounting for differences in transcript levels, as described by (Chothani et al., 2017), was used to compare SKIV2L binding for these two cell lines.

### Identification of rRNA binding sites by CRAC

The human ribosome structure was downloaded from PDB (4UG0) and its 18S rRNA sequence extracted. This was then substituted to match the mouse 18S rRNA sequence where possible, to facilitate alignment. CRAC reads with adapters and barcodes removed were aligned to this modified 18S rRNA sequence using bowtie2-2.2.3 (–sensitive mode). Alignments were filtered to remove those less than 20 bp long, or with an edit distance > 1. Reads were then piled up across the modified 18S rRNA sequence using samtools-1.3 (depth command) (Li et al., 2009). For each CRAC sample, these values were converted to single position counts per million, then mean normalized counts calculated for 20 nt bins across the entire 18S rRNA. This was repeated for 99 CRAC datasets, including several replicates each for SKIV2L, AVEN and XRN1, and an in-house collection of datasets from many unrelated proteins or untagged cell lines which served as controls. This extensive control dataset enabled specific signal to be robustly distinguished from background or technical artifacts. We also included additional SKIV2L, AVEN and XRN1 datasets for which global coverage was too low for mRNA analysis, but rRNA coverage sufficiently high for rRNA analysis. All raw data are deposited in GEO under accession GSE134020 and the control datasets indicated.

To quantify specific binding of AVEN, a negative binomial model was then used to fit the AVEN and control values for each 20 nt 18S rRNA bin. This model contained AVEN versus control as a factor, and was compared to a null (intercept only) model using a $\chi 2$ test (accounting for multiple hypothesis testing using the Benjamini Hochberg method with an FDR of 0.05). The $\chi 2$ p values were then used to color significantly bound regions of the 18S rRNA, in the context of the 40S ribosome structure. This procedure was repeated for SKIV2L and XRN1. Note that SKIV2L and XRN1 were included as controls for each other, and XRN1 (but not SKIV2L) was included as a control for AVEN.

### CRAC plots around start and stop codons

Uniquely mapping CRAC reads were piled up across each Appris transcript, and each transcript normalized by dividing by its the maximum read depth, and transcripts with fewer than five reads excluded. These normalized values were then either plotted around the start or stop codon of each individual transcript (arranging transcripts by 5′ UTR or 3′ UTR length), or values were summed to produce a metaplot.

### CRAC enrichment at amino acid combinations

The 3′ end positions of filtered uniquely mapping (but not low complexity stripped) SKIV2L CRAC reads were extracted for all mRNAs. These reads were further filtered to retain only those for which the 3′ adaptor had been identified and removed (so the 3′ end of the remaining read corresponds to the true 3′ end of a captured RNA fragment). Taking mRNAs with at least five CRAC reads passing these filters, 204 nt sliding windows were generated across the CDS, with an offset of 1 between each window. For each window, the hexamer (6 nt sequence) at its center was recorded, together with its frame (0, 1 or 2) relative to the start codon. SKIV2L CRAC 3′ ends were then piled up across every window for each transcript, and windows with at least five counts retained. The values for each position of a given window were then divided by the sum for that window. Windows were then grouped by their central hexamer (e.g., AAAAAG) and its frame (e.g., 0), and pileups summed for each group, dividing by the number windows within each

group. This provides an average distribution of SKIV2L-bound RNA fragment 3′ ends around every 6 nt motif, for frames 0, 1 and 2. This process was repeated for XRN1, RPS10 (a ribosomal protein) and TRIM71 CRAC datasets, which were used as controls.

For each dataset, hexamer and frame, the values for the central 6 nt were summed (these correspond precisely to the hexamer). ''SKIV2L-bound'' hexamer/frame combinations were then defined as those for which the SKIV2L value was higher than any of the other (control) datasets, and higher than that expected if SKIV2L reads had been distributed evenly across the 204 nt window. Hexamers were then translated into amino acid pairs (e.g., AAAAAG = > KK), then for each amino acid pair, the proportion of SKIV2L-bound hexamers calculated. This was repeated for in-frame hexamers (frame 0) and out of frame hexamers (frames 1 and 2). Only amino acid pairs with at least four contributing hexamers were evaluated. These final values (ranging from 0 to 1) give an indication of whether SKIV2L binds preferentially to particularly amino acid pairs, and whether this is frame specific (i.e., likely to reflect the encoded amino acids) or not (i.e., likely to reflect the underlying sequence).

### CRAC and disome profiling repeat analyis

For plots around amino acids repeats, the 3′ ends of uniquely mapping CRAC reads (not low complexity stripped) for which adapters were detected and removed, and uniquely mapping ribosome profiling reads, were used. These were piled up across windows centered on all 3-4 amino acid repeat tracts in Appris mRNAs (e.g., KKKK, AAA, etc), including 96 nt either side. Data were binned into 6 nt bins and normalized to the maximum count for each window. Values were then summed for each repeat type (e.g., K, A, etc) at each position.

For plots around polypurine ($[G/A]_{12+}$) tracts, a similar approach was used, piling up CRAC 3′ read positions or ribosome profiling reads across windows centered on the GA tract and including 96 nt flanks. For each window, pileups were normalized to the total counts, then pileups summed for windows grouped by G or A content, or grouped by the encoded amino acids (e.g., those with > 30% lysine). Note that repeat tracts were only identified in-frame, and were required to be a multiple of three nt long (to enable them to be translated into an amino acid sequence).

### RNA-seq analysis

RNA-seq reads were aligned to mm10 and counted using STAR_2.5.0a (–runMode alignReads –outSAMtype BAM SortedByCoordinate–outFilterType BySJout–outFilterMultimapNmax 1–outFilterMismatchNmax 3–outSAMmultNmax 1–out SAMattributes NH HI NM MD AS nM–outMultimapperOrder Random–outSAMunmapped None–quantMode GeneCounts). The STAR index was made using the Mus_musculus.GRCm38.75.dna.primary_assembly.fa file, providing gencode.vM16.annotation.gtf as the sjdbGTF file. DESeq2 (Love et al., 2014) was used to test for differential gene expression, with the model formula including biological clone (where at least two were available), sequencing batch (where more than one was performed) and genotype/treatment (WT versus knockout, or time of 4OHT treatment).

### RNA half-life analysis

Mapped RNA-seq reads (ERCC sequences were included as extra chromosomes for the mapping) were counted for each gene and ERCC using the–quantMode GeneCounts mode in STAR. Counts were normalized separately for each time point using the estimateSizeFactorsForMatrix function from DESeq2, then recombined. A single size factor was calculated for each time point to account for the overall decay of mRNAs. For this, the ratio of total mRNA counts to total ERCC counts (using the set of ERCCs with a mean of > 50 counts across all samples) was calculated for each sample. A median value (size factor) was then calculated for each time point, size factors scaled so that the size factor for time = 0 was 1, then all mRNA count tables for a given time point divided by the corresponding size factor.

Half-lives were then calculated by fitting a linear model for ln(normalized counts + 1) versus time, and using the formula $t_{1/2} = -\ln(2)/k$, where k is the coefficient for time (i.e., the gradient of the linear fit in semi-log space). The residual standard error was also calculated, as a measure of fit.

### CRAC sequence and structure motif analysis

Our approach was based on that described in (Welte et al., 2019). Filtered CRAC reads aligned to mRNAs robustly detected in ribosome profiling experiments were counted for 50 nt sliding windows (offset 10 nt) across the 5′ UTRs and CDSes of these transcripts. These windows were also folded *in silico* using RNAfold 2.1.5 (Lorenz et al., 2011) (including the option -g, to incorporate G-quadruplex formation into the prediction). A given ''foreground'' sample (e.g., AVEN) together with several control datasets (''background'' samples, comprising MTR4, XRN1, RPS10 (a ribosomal protein) and TRIM71 (Welte et al., 2019) were analyzed, and for each dataset, counts converted to RPKM. Median RNA-seq RPKM values for our SKIV2L, XRN1 and MTR4 tagged mESCs were also extracted for each transcript, removing transcripts with < 10 RPKM. CRAC values were then divided by RNA-seq values to obtain enrichments for each window and foreground or background dataset. ''Bound'' windows were defined as those where the foreground enrichment was higher than any of the background enrichments, and at least 10 (i.e., 10x CRAC coverage versus RNA-seq coverage). The highest bound 5′ UTR and CDS window was then selected for each transcript (final foreground window set). Transcripts with fewer than six analyzed windows were excluded. As a control, windows were randomly selected from the same transcripts, requiring them to be within 400 nt of the foreground window set, and excluding the foreground window set. This process was repeated 100 times, to generate a 100 final background window sets.

For the final foreground and each final background window set, the total number of windows containing each possible 10-mer structural motif (based on the RNAfold output) or 6-mer sequence motif were counted. For each motif, the mean background total and its standard deviation were calculated, and used to calculate a z-score (foreground minus mean background occurrence, divided by background standard deviation). The z-score was plotted for each motif, versus its $\log_2$ total occurrence (foreground plus mean background). Motifs with a z-score magnitude > 2.5, and sufficiently high $\log_2$ total occurrence, were highlighted.

To compare SKIV2L CRAC data from WT and *Aven*$^{-/-}$ cells, the same approach was used, except both rather than comparing the number of bound windows containing each structure/sequence motif to a randomly sampled set of windows, SKIV2L-bound windows for the two cell lines were compared directly.

### CRAC/ribosome profiling at structured regions

For mRNAs robustly detected in ribosome profiling experiments, 50 nt non-overlapping windows were defined across the 5′ UTR and CDS, and folded *in silico* using RNAfold 2.1.5. Windows with a minimum free energy $< -12$ kcal/mol and continuous stretch of $\geq 10$ paired nucleotides were selected, and extended by 96 nt either side. CRAC and ribosome profiling reads were piled up across these 242 nt windows, and these values normalized for each dataset and window to the window sum. Windows were then divided into 6 nt bins, and normalized counts summed for each bin and plotted.

### CRAC/ribosome profiling for genomic windows

For this analysis, CRAC reads uniquely mapping to the genome were used. RNA-seq reads, and ribosome profiling reads that had been trimmed, quality filtered, size selected ([26,35] for monosome footprints, [45,70] for disome footprints and [21,70] for total input RNA), and filtered against rRNA and tRNA libraries, were mapped to the genome using STAR_2.5.0a (settings–runMode alignReads–outSAMtype BAM SortedByCoordinate–outFilterType BySJout–outFilterMultimapNmax 1–outFilterMismatchNmax 3–outSAMmultNmax 1–outSAMattributes NH HI NM MD AS nM–outMultimapperOrder Random–outSAMunmapped None–quantMode GeneCounts. CRAC, RNA-seq and ribosome profiling reads were then counted (in a strand specific manner) for 1 kb windows tiling the genome in both orientations, generated using bedtools (Quinlan, 2014).

Windows were also overlapped with protein-coding genes, or genes encoding abundant ncRNAs (e.g., rRNA, snRNA and snoRNA genes) but not lincRNAs. This enabled windows to be classified based upon whether they overlapped abundant ncRNAs, mRNA exons (sense orientation), mRNA introns (sense orientation), 1 kb regions upstream of mRNAs (sense orientation), mRNA exons (antisense orientation), mRNA introns (antisense orientation) or 1 kb regions upstream of mRNAs (antisense orientation), with priority given in that order (i.e., if a window overlapped an mRNA exon and a ncRNA, it would be classified as a ncRNA window). All other windows were classified as intergenic. GC contents, predicted minimum free energy (using RNAfold), potential small ORFs (sequences starting with ATG and ending with an in-frame stop codon), and average phyloCSF scores for these small ORFs, were also calculated for each window.

Window counts were then normalized using the DESeq2 function "estimateSizeFactorsForMatrix" (for RNA-seq and ribosome profiling) or to the minimum library size (for CRAC). Normalized counts were then $\log_2$ transformed, including a pseudocount of 4. $\text{Log}_2$ fold changes were then calculated for *Aven*$^{-/-}$ versus WT datasets, or comparing CRAC datasets (e.g., AVEN versus SKIV2L), as indicated. Genomic windows were also classified based on their SKIV2L CRAC counts in *Aven*$^{-/-}$ versus WT, into the categories "down," "slight change," "up" and "strongly up." This enabled $\log_2$ fold changes (e.g., comparing RNA-seq for *Aven*$^{-/-}$ versus WT) to be median centered on the "down" category of windows, facilitating comparison of different data types. Where more than one batch was available for RNA-seq datasets, batches were merged by calculating mean values for each window at the end of the above procedure.

### Ribosome profiling analysis

#### *Preprocessing of Ribosome Footprints*

Initial quality assessment of the sequencing reads was conducted based on the preliminary quality values produced by the Illumina pipeline 2.19.1 such as the percentage of clusters passed filtering (%PF clusters) and the mean quality score (PF clusters). Adaptor sequences were removed using cutadapt utility (Martin, 2011) with following options: -a AGATCGGAAGAGCACACGTCT GAACTCCAGTCAC–match-read-wildcards. Next, trimmed read sequences were filtered by their size using an in-house Python script with following inclusive ranges: [26,35] for monosome footprints, [45,70] for disome footprints and [21,70] for total RNA. Smaller or larger fragments were kept separately and not used in further analysis. Finally, the reads were filtered for quality using fastq_quality_filter tool from the FASTX-toolkit with the following arguments: -Q33 -q 30 -p 80.

#### *Mapping of Footprints to Mouse Genome*

The preprocessed insert sequences were mapped sequentially to following databases: mouse rRNA, human rRNA, mt-tRNA, mouse tRNA, mouse cDNA from Ensembl mouse database release 91 (Flicek et al., 2013) and, finally, mouse genomic sequences (Genome Reference Consortium GRCm38.p2). In all but the final mapping against genomic sequences, bowtie version 2.3.0 (Langmead and Salzberg, 2012) was used with the following parameters: -p 2 -L 15 -k 20–no-unal,

After each alignment, only reads that were not aligned were used in the following mapping. For further analysis, only alignments against mouse cDNA were used, unless specifically stated otherwise. For each query sequence, only alignments with maximum alignment scores were kept.

Separately from this sequential alignment strategy, trimmed and filtered total RNA sequences from each sample were also directly aligned against the mouse genome. This mapping and the final mapping of the sequential alignment strategy were performed using STAR version 2.5.3a (Dobin et al., 2013) with the following parameters:

```
--runThreadN 6 –genomeDir=mouse/star/Mmusculus.GRCm38.91
--readFilesCommand zcat –genomeLoad LoadAndKeep
--outSAMtype BAM SortedByCoordinate Unsorted
--alignSJDBoverhangMin 1 –alignIntronMax 1000000
--outFilterType BySJout –alignSJoverhangMin 8
--limitBAMsortRAM 15000000000
```

The output of this alignment was used to estimate expressed transcript models out of all models contained in Ensembl mouse database release 91. To this end, we used StringTie version 1.3.3b (Pertea et al., 2015) to estimate the number of fragments per kilo base of exon per million fragments mapped (FPKM) for each transcript, with the following parameters:

```
-p 8 -G Mmusculus.GRCm38.91.gtf -A gene_abund.tab
-C cov_refs.gtf -B -e
```

The resulting FPKM estimate information was parsed with an in-house Python script to identify transcripts which had an FPKM > 0.2 and an isoform abundance fraction > 0.05 in at least 2 samples. A database of expressed transcripts based on this filtering was used in further analysis. Among those, genes that were estimated to have a single expressed isoform were annotated as single transcript genes.

### *Quantification of Footprint Abundance/Density*

Abundance of mRNA and monosome or disome protected fragments was estimated per gene as described in (Janich et al., 2015). For this quantification, only reads that were mapped uniquely to a single gene and only to transcripts that were identified to be expressed (see Mapping of Footprints to Mouse Genome) were used. We used a limited size range of disome fragments (56-64 nt), as this facilitated subsequent A-site assignment and high-resolution analysis of stall sites.

Read counts of total RNA and RPF were normalized with upper quantile method of R package edgeR v3.16.5 (Robinson et al., 2010). Prior to normalization, transcripts which did not have at least 10 counts in at least one third of the samples were removed from the datasets. For better comparability between datasets, RPKM values were calculated as the number of counted reads per 1000 mappable and countable bases per geometric mean of normalized read counts per million. Genes that had an average total RNA RPKM > 5 were designated as robustly expressed.

Ribosome densities (alternatively known as translational efficiencies, TE) were then calculated as the ratio of footprint-RPKM to total RNA-RPKM for monosomes and disomes per sample. For most analysis downstream, densities were $\log_2$ transformed. Significant changes in abundances of total RNA, monosomes and disomes between control and treated samples were assessed by DESeq2 package for the R statistical environment (Team, 2013). Significant changes in densities of monosomes and disomes between control and treated samples were assessed by R package xtail. The false discovery rate (FDR) adjusted p values were used to identify statistically significant changes at 0.05 FDR.

### *Monosome and Disome Positions on Transcripts*

For total RNA and RPF reads that were counted toward genes, we have also tracked the position of the reads relative to the 5′ end of its corresponding transcript. For total RNA reads and monosome footprints we have used the 5′ end of the reads and the estimated A-site of the ribosomes, respectively, as described in (Janich et al., 2015). For disome footprints we have established an empirical offsetting scheme based on the size and the frame (relative to the main CDS) of the footprints. We estimated the A-site of the pausing ribosome at the disome site by adding 45, 44, or 43 to the map position of the 5′ end of 58nt-long foot- prints that were at 1st, 2nd or 3rd frame, respectively. Similarly, we used the following offsets for 59nt, 60nt, 62nt and 63nt long disome footprints, respectively: [45, 44, 46], [45, 44, 46], [48, 47, 46], [48, 47, 49]. These coordinates then were converted into Wiggle Track Format (WIG) by in-house Python scripts.

### *Ribosome profiling versus CRAC for uORFs*

uORFs were defined using our ribosome profiling data, as described in (Castelo-Szekely et al., 2019). Briefly, transcripts that are the only protein-coding isoform expressed were used (n = 7593), so that footprints can be unambiguously assigned to the 5′ UTR. uORFs were annotated based on the following criteria: 1) started with AUG, 2) had an in-frame stop codon within the 5′ UTR or within the CDS (overlapping uORFs) and 3) were at least 6 nt long (including the stop codon).

CRAC and ribosome profiling reads overlapping with uORFs and main CDSes were counted, and normalized to 100000 for each sample. To group uORFs by AVEN occupancy in WT cells, AVEN uORF CRAC was normalized to uORF ribosome profiling counts in WT cells, then uORFs classified by this value (low $\leq$ 5; medium > 5 and $\leq$ 20; high > 20). Differential uORF translation for *Aven*$^{-/-}$ versus WT cells was then calculated by normalizing uORF ribosome profiling counts to main CDS ribosome profiling counts for the two cell lines, then calculating a $\log_2$ fold change.

A similar approach was used to compare SKIV2L binding to uORFs in *Aven*$^{-/-}$ versus WT cells, whereby SKIV2L uORF CRAC counts were normalized to main CDS ribosome profiling counts (to account for changes in overall mRNA translation) before

calculating a $\log_2$ fold change for *Aven*$^{-/-}$ versus WT. Only uORFs with at least 20 counts per 100k for either SKIV2L in WT, or SKIV2L in *Aven*$^{-/-}$, were used. Additionally, the main CDS was required to have at least 2 counts per 100k for monosomes in both WT and *Aven*$^{-/-}$ conditions. The rationale here was to only include uORFs where uORF:CDS ratios could be calculated without being dominated by noise/background.

## DATA AND CODE AVAILABILITY

The accession number for the sequencing data reported in this paper is GEO: GSE134020. R code and scripts used for analysis are available upon request. Original western blots were deposited in Mendeley Data and are available at https://data.mendeley.com/datasets/c6zdfw957p/1

**Molecular Cell, Volume *77***

## Supplemental Information

# Mammalian RNA Decay Pathways Are Highly Specialized and Widely Linked to Translation

**Alex Charles Tuck, Aneliya Rankova, Alaaddin Bulak Arpat, Luz Angelica Liechti, Daniel Hess, Vytautas Iesmantavicius, Violeta Castelo-Szekely, David Gatfield, and Marc Bühler**

## SUPPLEMENTAL FIGURE LEGENDS

**Figure S1: Profiling mammalian RNA decay pathways (related to Figure 1). A** Western blots showing expression and biotinylation of endogenously tagged 3xFLAG-AviTag SKIV2L, XRN1 and MTR4. **B** Scree plot for the PCA in Figure 1C. **C** Western blot analysis of a co-immunoprecipitation using anti-HA to capture 2xHA-FKBP12-tagged DIS3L, and probing inputs and eluates with anti-FLAG (to detect 3xFLAG-Avi-SKIV2L) and anti-MTR4. **D** tSNE representation of mRNAs based on relative binding to MTR4, SKIV2L and XRN1, and highlighting functional classes of mRNAs. **E** RT-qPCR analysis of *Skiv2l*$^{fl/fl}$ mRNA expression levels in clones 5F and 9E (Table S1) after 0.1 μM 4OHT treatment for 0, 2, 4 and 6 days. Error bars denote standard deviation of two technical replicates. Values are normalized to TBP expression and then to untreated samples (day 0). **F** Boxplot for the 7240 mRNAs analyzed in Figure 1I, grouped by differential expression after four days of *Skiv2l* knockout, and showing the extent to which they bind SKIV2L in WT cells (relative binding based on CRAC, defined in Figure 1E). Box widths are proportional to the number of contained transcripts. $p < 10^{-15}$ for upregulated versus downregulated transcripts (Mann-Whitney U test). **G** $Log_2$-fold changes in transcriptome-wide mRNA half-lives for *Skiv2l*$^{fl/fl}$ cells following 0.1 μM 4OHT treatment for 4 or 0 days. Half-lives were calculated by fitting an exponential decay model to RNA-seq counts from an actinomycin D-mediated transcription shut-off time course. The x-axis shows the extent of SKIV2L binding as a fraction of total SKIV2L+XRN1 binding (i.e. SKIV2L CRAC divided by SKIV2L+XRN1 CRAC, using "relative binding" values defined in Figure 1E). High confidence SKIV2L targets are indicated in red, and are defined as mRNAs with an increased half-life after *Skiv2l* knockout (y-axis > 0.1), high SKIV2L binding (x-axis > 0.5) and significant RNA increase after 4 days of *Skiv2l* knockout (Figure 1I). The half-lives of the 200 most highly SKIV2L-bound mRNAs (right of green dashed line) were compared to those of all other mRNAs using a Student's t-test. **H** RNA levels for high confidence SKIV2L targets (red circles in G) for wild-type, tagged (*Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$, *Mtr4*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ and *Xrn1*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$) and *Skiv2l*$^{fl/fl}$ (+ 4 days 4OHT) cells. Each point refers to a separate cell line. See also Tables S1-S3.

**Figure S2: Defining triggers of RNA decay (related to Figure 2). A** Proportion of CRAC mRNA reads with non-templated homopolymeric 3' tails, for XRN1, SKIV2L and MTR4. **B** SKIV2L specific enrichment at amino acid pairs, compared to a panel of control datasets. For each amino acid pair, the proportion of hexamers encoding that amino acid pair and bound by SKIV2L is shown. This calculation was performed for in-frame hexamers (y-axis) and out-of-frame hexamers (x-axis). **C** Disome profiling (A sites) and SKIV2L and XRN1 CRAC coverage (3' end of RNA fragments) around runs of 3-4 identical amino acids (e.g. KKK/KKKK), normalized to downstream signal. **D** Pileups around polypurine tracts for monosome and disome profiling and SKIV2L, MTR4 and XRN1 CRAC signal. Polypurine tracts are divided by lysine content (top) or A versus G content (≥ 50 % A, or > 50 % G).

**Figure S3: AVEN modulates SKIV2L binding (related to Figure 4). A** Scree plot for the PCA in Figure 4C. **B** Left, schematic of CRISPR-Cas9-induced *Aven* knockout and RNA-seq tracks in the parental *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ and *Aven*$^{-/-}$ clones 4H and 6G. Right, western blot analysis of AVEN expression. **C** RT-qPCR analysis of *Focad* mRNA levels in *Focad*$^{-/-}$ clones 2F and 4B and WT cells. Error bars denote standard deviation of four technical replicates. **D and E** Boxplot representation of the data shown in Figure 4F, including all transcripts (not only those differentially bound or expressed). AVEN CRAC counts in WT are shown (y-axis) for transcripts grouped by changes in SKIV2L binding (B; $p < 10^{-15}$) or RNA abundance (C; $p = 0.002$) in *Aven*$^{-/-}$ versus WT (x-axis). P-values calculated using a Mann-Whitney U test to compare up- and down-regulated transcripts. See also Tables S1 and S3.

**Figure S4: AVEN and SKIV2L counteract ribosome stalling (related to Figures 4 and 5). A** Changes in mRNA monosome and disome densities in *Aven*$^{-/-}$ versus WT. Transcripts are colored by AVEN binding in WT cells (calculated as in Figure 4F; high, n = 777; low, n = 5471), and a linear best fit line plotted for each group (shaded area = 95 % confidence interval). **B** Western blot analysis of SKIV2L levels in *Aven*$^{-/-}$

*Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Skiv2l*$^{fl/fl}$ cells after treatment with 0.1 μM 4OHT for 0, 2 or 4 days to induce *Skiv2l* knockout. **C** $Log_2$ fold change in the half-lives of AVEN targets, which are defined as mRNAs that accumulate upon combined knockout of *Aven* and *Skiv2l* ($p < 0.01$; Figure 5D) and that have increased SKIV2L binding in *Aven*$^{-/-}$ cells (differential binding > 0.5, Figure 5D). Three comparisons are shown (left to right): (i) *Skiv2l*$^{fl/fl}$ 0 vs 4 days 4OHT (i.e. *Skiv2l* knockout vs wild-type), (ii) *Skiv2l*$^{fl/fl}$ vs *Aven*$^{-/-}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Skiv2l*$^{fl/fl}$ with no 4OHT treatment (i.e. *Aven* knockout vs wild-type), and (iii) *Aven*$^{-/-}$ *Skiv2l*$^{3xFLAG\text{-}Avi/3xFLAG\text{-}Avi}$ *Skiv2l*$^{fl/fl}$ 0 vs 4 days 4OHT (i.e. *Skiv2l Aven* double knockout versus *Aven* knockout). Half-lives are calculated from RNA-seq decay curves following actinomycin D-mediated transcription shut-off. Replication-dependent histone mRNAs are colored red. For each of the three comparisons, a Student's t-test was used to test whether the distribution of half-life $log_2$-fold changes for the set of mRNAs shown here (AVEN targets) differs from that of all other mRNAs. **D** Top, scheme depicting mESC cell cycle synchronization at the G1/S boundary using a double-thymidine block. Bottom, flow cytometry analysis of DAPI-stained asynchronous or thymidine-synchronized cells at 0, 4 and 8 hours after release from the second thymidine block for the indicated cell lines. **E** Boxplot showing the change in uORF monosome profiling counts (normalized to main CDS counts) for *Aven*$^{-/-}$ versus WT cells. uORFs are categorized (x-axis) by their AVEN CRAC counts in WT cells (n = 558, 215 and 53 for low, medium and high categories; Mann-Whitney U test $p = 3.19 \times 10^{-6}$ comparing high and low categories). **E** Western blot analysis of ATF4 and eIF2α-phospho-Serine51 levels in WT and *Aven*$^{-/-}$ cells after 0, 0.5, 2 and 4 hours of treatment with 200 nM thapsigargin. See also Tables S3 and S4.

**Figure S5: Defining sites dependent on AVEN and SKIV2L (related to Figure 6). A** $Log_2$ fold changes in SKIV2L CRAC counts for 1 kb windows tiling the genome, in *Aven*$^{-/-}$ versus WT cells. Points are colored by the proportion of U-tailed reads within each window in WT (left) or *Aven*$^{-/-}$ (right) cells. **B** SKIV2L CRAC, disome and monosome profiling reads piled up around predicted structured regions (minimum free

energy <-12 kcal/mol and continuous stretch of paired nucleotides ≥10), normalized to the downstream region. For SKIV2L CRAC in *Aven*$^{-/-}$, U-tailed reads are also shown.

**Figure S6: Features of translated non-coding RNAs (related to Figure 7). A** GC content and **B** predicted free energy for the "non-coding" 1 kb genomic windows analyzed in Figure 7C. Windows are split into four categories based on differential SKIV2L CRAC counts in *Aven*$^{-/-}$ versus WT cells (defined in Figure 7C). **C** PhyloCSF scores for predicted small ORFs within these windows. Windows are classified based upon whether SKIV2L CRAC and monosome profiling counts increase in *Aven*$^{-/-}$ versus WT cells, or not. See also Tables S6 and S7.

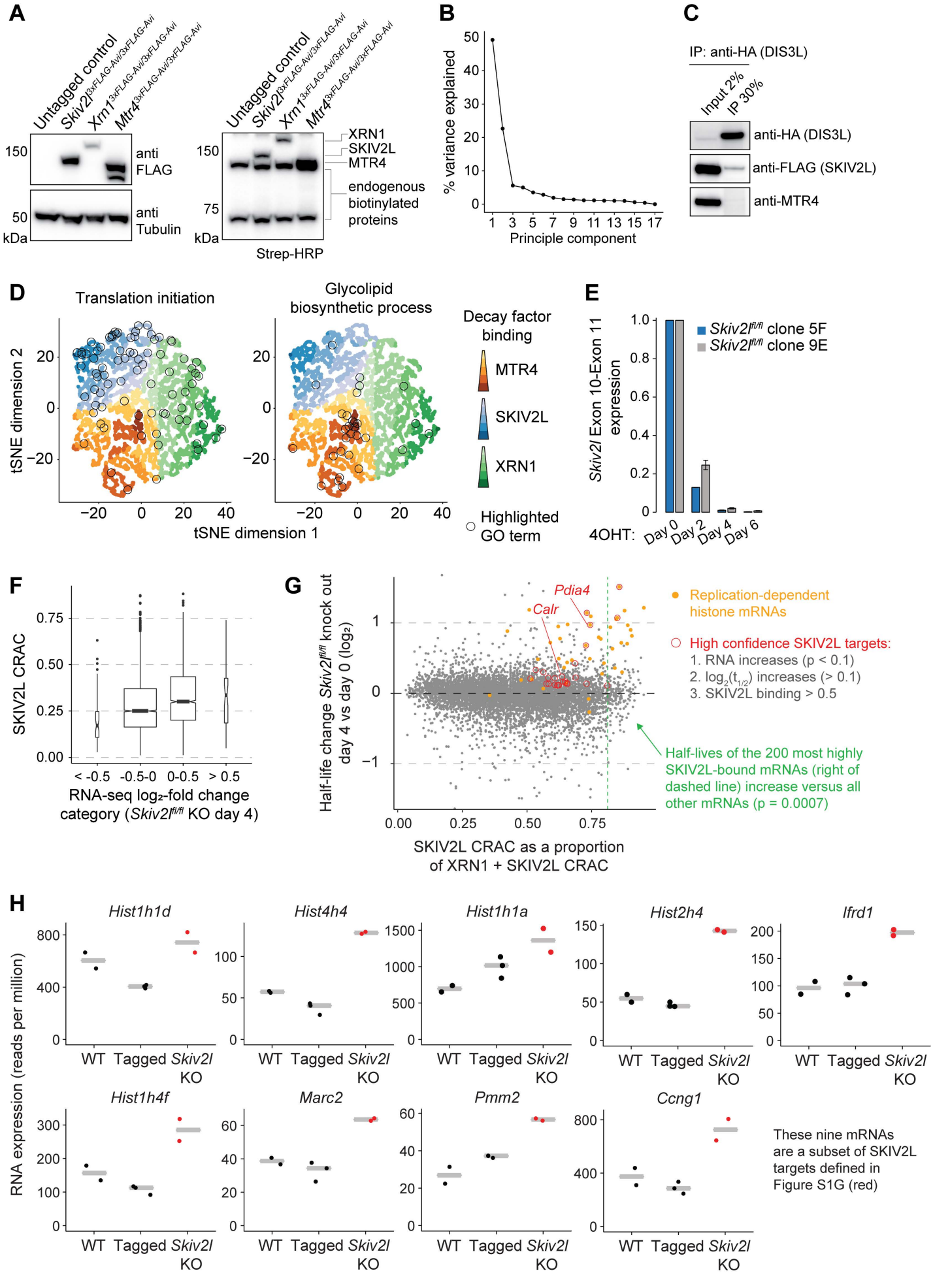

A
Untagged control
Skiv2l3xFLAG-Avi/3xFLAG-Avi
Xrn13xFLAG-Avi/3xFLAG-Avi
Mtr43xFLAG-Avi/3xFLAG-Avi
150
anti FLAG
50
anti Tubulin
kDa
XRN1
SKIV2L
MTR4
endogenous biotinylated proteins
75
Strep-HRP
B
% variance explained
Principle component
C
IP: anti-HA (DIS3L)
Input 2%
IP 30%
anti-HA (DIS3L)
anti-FLAG (SKIV2L)
anti-MTR4
D
Translation initiation
Glycolipid biosynthetic process
tSNE dimension 2
tSNE dimension 1
Decay factor binding
MTR4
SKIV2L
XRN1
Highlighted GO term
E
Skiv2l Exon 10-Exon 11 expression
Skiv2lfl/fl clone 5F
Skiv2lfl/fl clone 9E
4OHT:
Day 0
Day 2
Day 4
Day 6
F
SKIV2L CRAC
RNA-seq log2-fold change category (Skiv2lfl/fl KO day 4)
< -0.5
-0.5-0
0-0.5
> 0.5
G
Half-life change Skiv2lfl/fl knock out day 4 vs day 0 (log2)
SKIV2L CRAC as a proportion of XRN1 + SKIV2L CRAC
Pdia4
Calr
Replication-dependent histone mRNAs
High confidence SKIV2L targets:
1. RNA increases (p < 0.1)
2. log2(t1/2) increases (> 0.1)
3. SKIV2L binding > 0.5
Half-lives of the 200 most highly SKIV2L-bound mRNAs (right of dashed line) increase versus all other mRNAs (p = 0.0007)
H
RNA expression (reads per million)
Hist1h1d
Hist4h4
Hist1h1a
Hist2h4
Ifrd1
Hist1h4f
Marc2
Pmm2
Ccng1
WT
Tagged Skiv2l KO
These nine mRNAs are a subset of SKIV2L targets defined in Figure S1G (red)

Figure S1

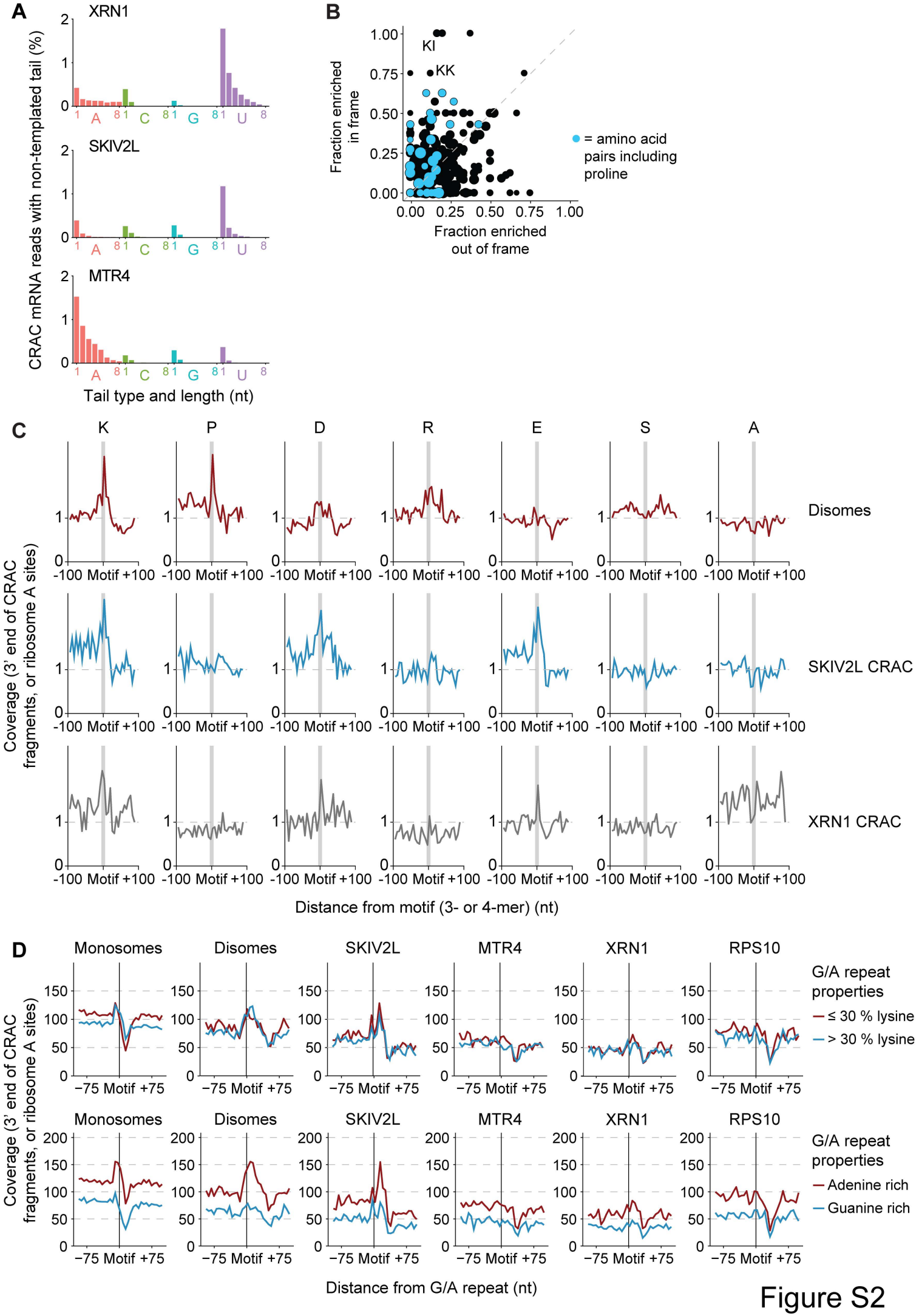

A
XRN1
SKIV2L
MTR4
CRAC mRNA reads with non-templated tail (%)
Tail type and length (nt)
B
Fraction enriched in frame
Fraction enriched out of frame
KI
KK
= amino acid pairs including proline
C
K
P
D
R
E
S
A
Disomes
SKIV2L CRAC
XRN1 CRAC
Coverage (3' end of CRAC fragments, or ribosome A sites)
-100 Motif +100
Distance from motif (3- or 4-mer) (nt)
D
Monosomes
Disomes
SKIV2L
MTR4
XRN1
RPS10
G/A repeat properties
≤ 30 % lysine
> 30 % lysine
Adenine rich
Guanine rich
−75 Motif +75
Distance from G/A repeat (nt)

Figure S2

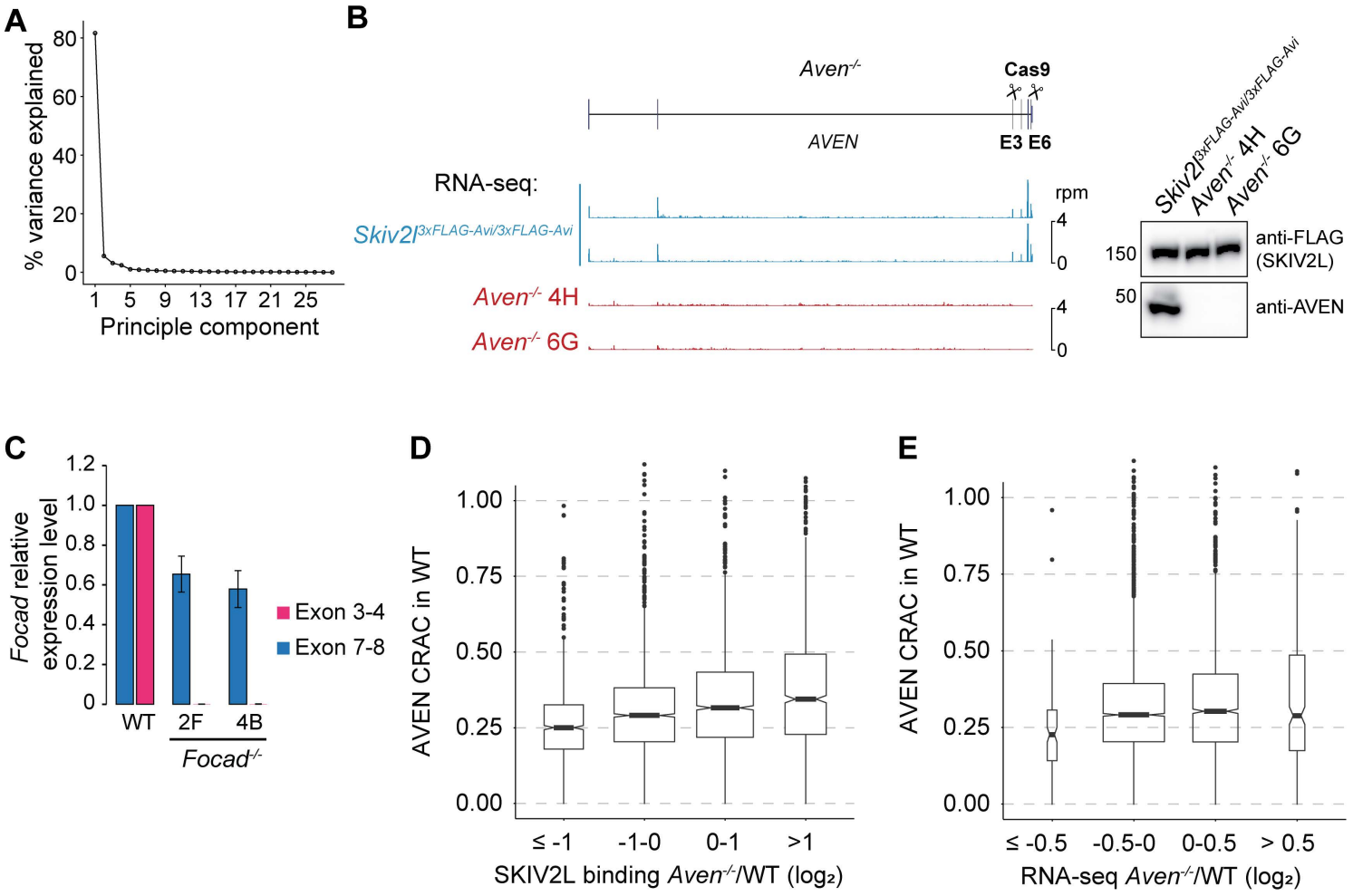
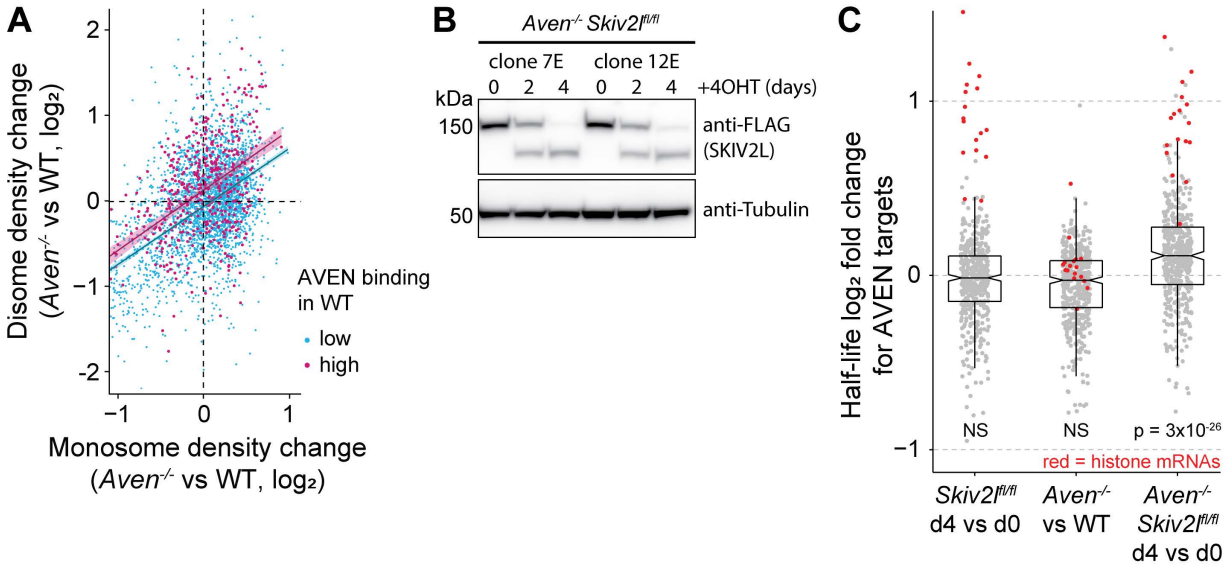
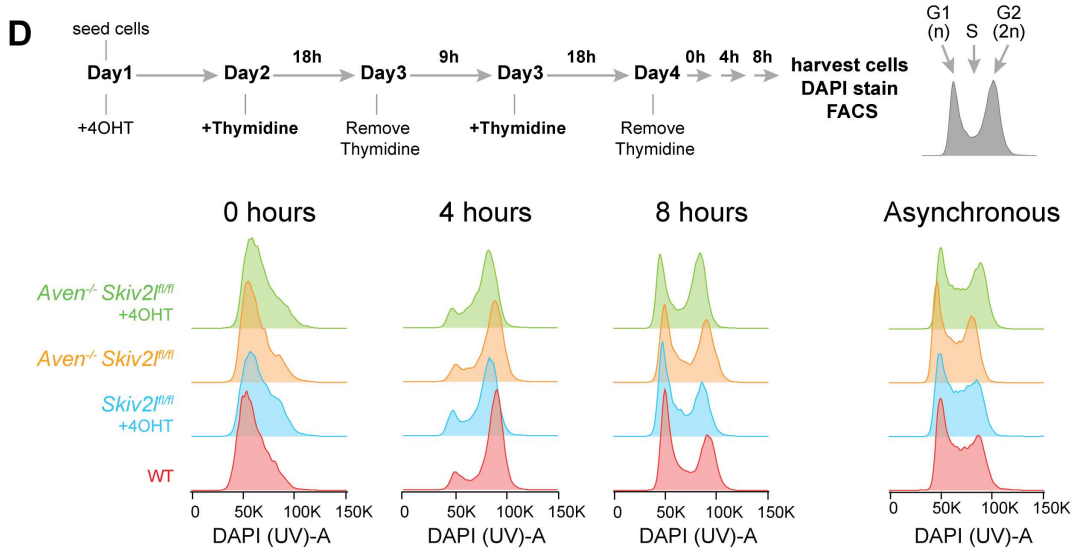
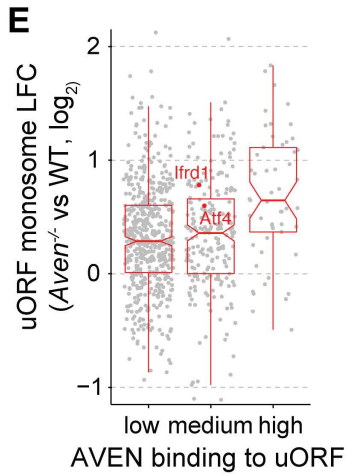

A
% variance explained
80
60
40
20
0
1 5 9 13 17 21 25
Principle component
B
Aven-/-
Cas9
AVEN
E3 E6
RNA-seq:
Skiv2l3xFLAG-Avi/3xFLAG-Avi
rpm
4
0
Aven-/- 4H
Aven-/- 6G
Skiv2l3xFLAG-Avi/3xFLAG-Avi
Aven-/- 4H
Aven-/- 6G
150
anti-FLAG (SKIV2L)
50
anti-AVEN
C
Focad relative expression level
1.2
1.0
0.8
0.6
0.4
0.2
0
WT
2F
4B
Focad-/-
Exon 3-4
Exon 7-8
D
AVEN CRAC in WT
1.00
0.75
0.50
0.25
0.00
≤ -1
-1-0
0-1
>1
SKIV2L binding Aven-/-/WT (log2)
E
AVEN CRAC in WT
1.00
0.75
0.50
0.25
0.00
≤ -0.5
-0.5-0
0-0.5
> 0.5
RNA-seq Aven-/-/WT (log2)

Figure S3


A
Disome density change
(Aven-/- vs WT, log2)
Monosome density change
(Aven-/- vs WT, log2)
AVEN binding
in WT
low
high
B
Aven-/- Skiv2lfl/fl
clone 7E
clone 12E
0 2 4 0 2 4 +4OHT (days)
kDa
150
anti-FLAG
(SKIV2L)
50
anti-Tubulin
C
Half-life log2 fold change
for AVEN targets
NS
NS
p = 3x10-26
red = histone mRNAs
Skiv2lfl/fl
d4 vs d0
Aven-/-
vs WT
Aven-/-
Skiv2lfl/fl
d4 vs d0


D
seed cells
Day1
+4OHT
Day2
18h
+Thymidine
Day3
Remove
Thymidine
9h
Day3
+Thymidine
18h
Day4
Remove
Thymidine
0h 4h 8h
harvest cells
DAPI stain
FACS
G1
(n)
S
G2
(2n)
0 hours
4 hours
8 hours
Asynchronous
Aven-/- Skiv2lfl/fl
+4OHT
Aven-/- Skiv2lfl/fl
Skiv2lfl/fl
+4OHT
WT
0 50K 100K 150K
DAPI (UV)-A


E
uORF monosome LFC
(Aven-/- vs WT, log2)
Ifrd1
Atf4
low medium high
AVEN binding to uORF

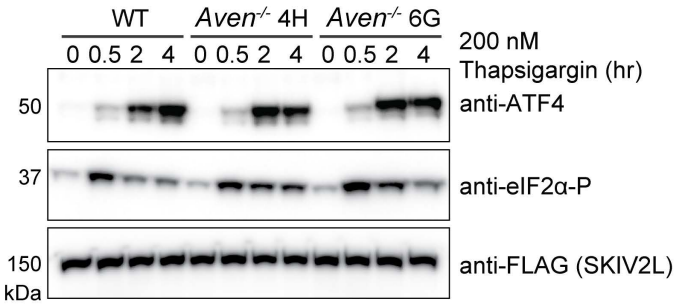

F
WT
Aven-/- 4H
Aven-/- 6G
0 0.5 2 4 0 0.5 2 4 0 0.5 2 4
200 nM
Thapsigargin (hr)
50
anti-ATF4
37
anti-eIF2α-P
150
kDa
anti-FLAG (SKIV2L)

Figure S4

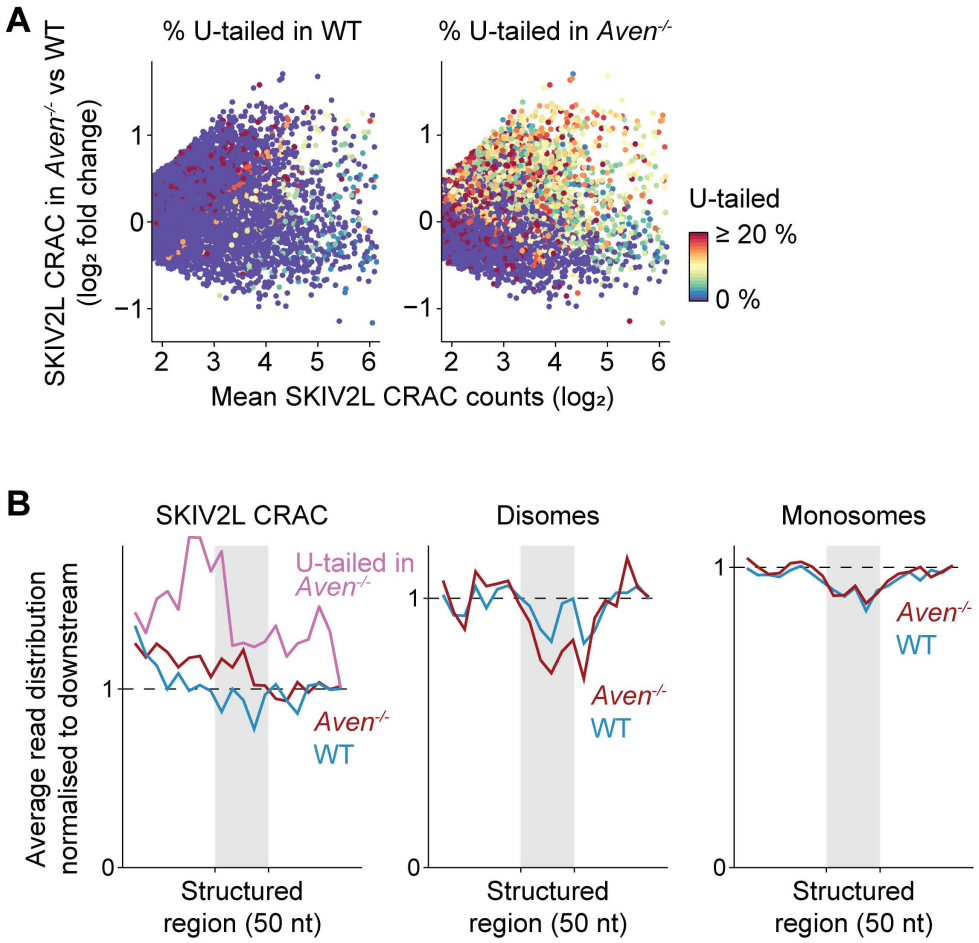

A
% U-tailed in WT
% U-tailed in Aven-/-
SKIV2L CRAC in Aven-/- vs WT
(log2 fold change)
1
0
−1
2
3
4
5
6
Mean SKIV2L CRAC counts (log2)
U-tailed
≥ 20 %
0 %
B
SKIV2L CRAC
Disomes
Monosomes
Average read distribution
normalised to downstream
U-tailed in
Aven-/-
Aven-/-
WT
Structured
region (50 nt)

Figure S5

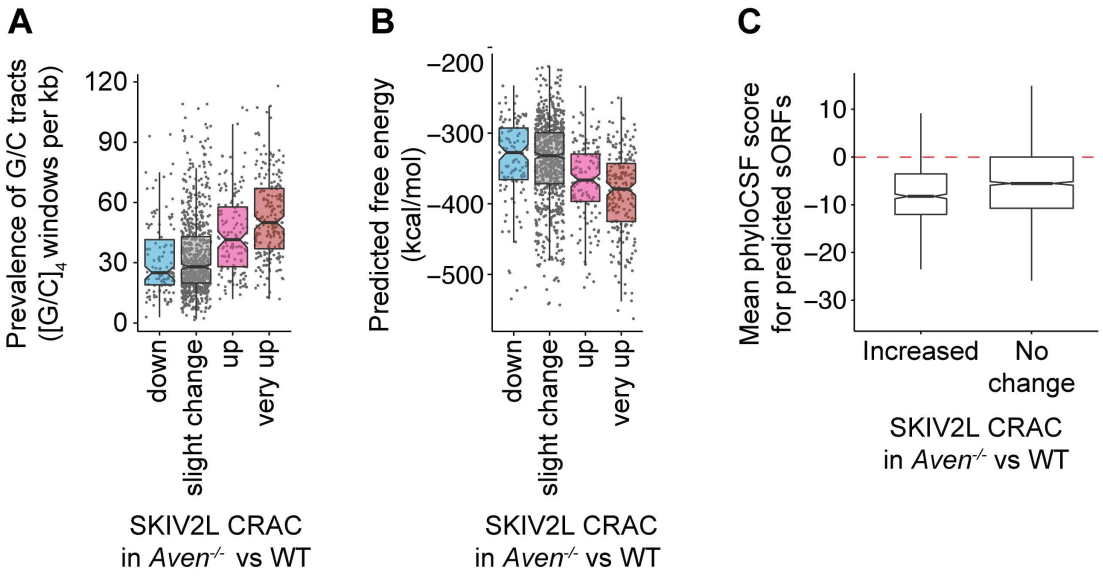

A
Prevalence of G/C tracts ([G/C]4 windows per kb)
120
90
60
30
0
down
slight change
up
very up
SKIV2L CRAC in Aven-/- vs WT
B
Predicted free energy (kcal/mol)
−200
−300
−400
−500
down
slight change
up
very up
SKIV2L CRAC in Aven-/- vs WT
C
Mean phyloCSF score for predicted sORFs
10
0
−10
−20
−30
Increased
No change
SKIV2L CRAC in Aven-/- vs WT

Figure S6